



(12)

## EUROPEAN PATENT APPLICATION

(21) Application number: 93105718.6

(51) Int. Cl.<sup>5</sup>: C12N 15/12, C12N 15/62,  
C07K 15/00, C12N 5/10

(22) Date of filing: 07.04.93

The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(30) Priority: 01.05.92 US 877675

(43) Date of publication of application:  
02.03.94 Bulletin 94/09

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB GR IE IT LI LU NL  
PT SE

(71) Applicant: AMERICAN CYANAMID COMPANY  
1937 West Main Street  
P.O. Box 60  
Stamford Connecticut 06904-0060(US)

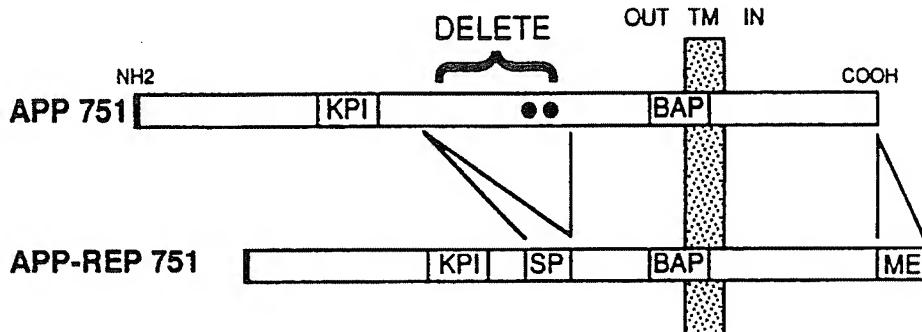
(72) Inventor: Vitek, Michael Peter  
213 Wyckoff Avenue  
Waldwick, New Jersey 07463(US)  
Inventor: Jacobsen, Jack Steven  
229 Mulberry Road  
Ramsey, New Jersey 07446(US)

(74) Representative: Wächtershäuser, Günter, Dr.  
Tal 29  
D-80331 München (DE)

(54) Novel amyloid precursor proteins and methods of using same.

(57) This invention provides novel nucleic acid molecules which encode amyloid precursor mutoins and the polypeptides encoded therefrom. Also provided are host vector systems useful for the recombinant production of the recombinant polypeptides in procaryotic and eucaryotic systems. Cells comprising the host vector systems of this invention as well as methods of recombinantly producing these polypeptides are provided by this invention. Further provided is a method to detect the recombinant polypeptides of this invention.

Figure 1.



BACKGROUND OF THE INVENTION

Throughout this application various references are referred to within parentheses. Disclosures of these publications in their entirety are hereby incorporated by reference into this application to more fully describe the state of the art to which this invention pertains. Full bibliographic citations for these references may be found at the end of this application, immediately preceding the claims.

Abnormal accumulation of extracellular amyloid in plaques and cerebrovascular deposits are characteristic in the brains of individuals suffering from Alzheimer's disease (AD) and Down's Syndrome (Glenner and Wong, BBRC, 120:885-890, 1984; Glenner & Wong, BBRC, 120:1131-1153, 1984). The amyloid deposited in these lesions, referred to as beta amyloid peptide (BAP), is a poorly soluble, self-aggregating, 39-42 amino acid (aa) protein which is derived via proteolytic cleavage from a larger amyloid precursor protein (APP) (Kang et al., Nature 325:733-736, 1987) BAP also is thought to be neurotoxic (Yankner et al., Science 245:417-420, 1990). APP is expressed as an integral transmembrane protein (Dyrks et al., EMBO J., 7:949-957, 1989) and is normally proteolytically cleaved by "secretase" (Sisodia et al., Science, 248:492-495, 1990; Esch et al., Science, 248:1122-1124) between BAP-16K (lysine) and - 17L (leucine). Cleavage at this site therefore precludes amyloidogenesis (Palmer et al., BBRC, 156:432-437, 1988) and results in release of the amino-terminal APP fragment which is secreted into tissue culture medium (Sisodia et al., ibid, Esch, et al., ibid). Three major isoforms of APP (APP-695, APP-751 and APP-770 amino acids) are derived by alternative splicing (Ponte, et al., Nature 331:525-527, 1988; Kitaguchi et al., Nature 331:530-532, 1988; and Tanzi, et al., Nature 331:528-530, 1988), are expressed as integral transmembrane proteins (Kang et al., Nature 325:733-736, 1987; Dyrks et al., EMBO J. 7:949-957, 1988).

Even though both APP-770 and -751 isoforms contain a protease inhibitor domain, it is the secreted portion of APP-751 (also known as Protease Nexin II (Van Nostrand et al., Science, 248:745-748, 1990) which is thought to be involved in cell adhesion (Schubert et al., Neuron, 3:689-694, 1989), remodeling during development, coagulation (Smith et al., Science, 248:1126-1128, 1990) and wound repair.

Although the mechanisms underlying abnormal proteolytic processes which result in BAP extraction from APP are poorly understood, it is thought to be central to the pathogenesis (Selkoe, Neuron, 6:487-498, 1991; Isiura, J. Neurochem. 56:363-369, 1991) and memory loss (Flood, et al., Proc. Natl. Acad. Sci. 88:3363-3366, 1991) associated with Alzheimer's Disease.

Based on the observations that (a) amyloid plaques develop in AD brains, (b) a major component of plaques is BAP, (c) BAP is generated by proteolytic cleavage of APP protein, (d) mRNA levels of specific APP isoforms increase in AD suggesting that more APP protein is expressed, (e) APP point mutations which are thought to possibly after normal processing have been identified in Familial AD (FAD) and "Dutch" disease, (f) injection of BAP into the brains of rodents both form lesions reminiscent of plaque pathology and result in memory deficits, and (g) the detection of plaque-like amyloid deposits in the brains of transgenic mice expressing human APP, it is important to understand how APP is processed to generate BAP.

SUMMARY OF THE INVENTION

This invention provides novel nucleic acid molecules which encode amyloid precursor mureins and the polypeptides encoded therefrom. Also provided are host vector systems useful for the recombinant production of the recombinant polypeptides in prokaryotic and eucaryotic systems. Cells comprising the host vector systems of this invention as well as methods of recombinantly producing these polypeptides are provided by this invention. Further provided is a method to detect the recombinant polypeptides of this invention.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: Schematic representation of APP-REP 751. APP-REP 751 represents a cleavable APP substrate system which contains target sequences of BAP including normal flanking regions (not to scale). The APP-REP protein is marked with a 276 amino acid deletion (corresponding to APP-751 beginning at Xhol through to and including the glycine codon at 15 amino acid residues N-terminal to BAP) and the insertion of sequences encoding N- and C-terminal reporter epitopes. Substrate P (SP) reporter epitope (RPKPQQFFGLM) is inserted at the Xhol site. Met-enkephaline (ME) reporter epitope (YGGFM) is inserted at the C-terminus of APP. The resulting construct encodes 492 amino acids (see Figure 2).

Figure 2: Schematic representation depicting the construction of APP-REP from APP-751 cDNA. Partial representing N- and C-terminal regions of APP-REP were cloned separately as illustrated below. The N-

terminal partial was constructed by ligating sequences encoding substance P (SP) to an N-terminal fragment of APP cDNA. The C-terminal partial was constructed by PCR amplification using the corresponding portion of APP cDNA to introduce novel ends including the Met-enkephalin (ME) reporter epitope. A functional APP-REP 751 clone was obtained by subcloning the partials as indicated. EcoRI (E), Xhol (X),  
5 HindIII (H), BamHI (B), Sall (S), XbaI (Xb).

Figure 3: Epitope mapping of APP-REP 751 expressed in COS-1 cells. Immunoprecipitation analysis of cell lysate and conditioned medium using the SP (anti-N-terminal substance P reporter) and M3 (anti-C-terminal APP) antisera. Lanes 1 and 2, cell lysate immunoprecipitated with SP and M3 antisera, respectively; lanes 3 and 4, conditioned medium immunoprecipitated with M3 and SP antisera, respectively; lanes 5 and 6, conditioned medium of control cells transfected with vector DNA immunoprecipitated with SP and M3 antisera, respectively; lane M, molecular weight markers.

Figure 4: Pulse-chase analysis of APP-REP 751. Immunoprecipitation of cell lysate (A) and CM (B). COS-1 cells were pulsed with [<sup>35</sup>S]-methionine for 15 minutes and chased using cold methionine for 0, 0.5, 1, 1.5, 2 and 4 hours (lanes 1 to 6). Lanes 7, 8 and 9 are chase intervals of 0, 1 and 2 hour for control cells transfecte with vector DNA. Lane M, molecular weight markers.

Figure 5: Epitope mapping and comparative expression of APP-REP 751, BAP<sub>E22Q</sub> and BAP<sub>Δ11-28</sub>. A, Schematic representation of relevant BAP (boxed) and flanking amino acid sequences of APP-REP 751, BAP<sub>E22Q</sub> and BAP<sub>Δ11-28</sub> juxtaposed against the putative transmembrane domain (shadowed). B-F, Immunoprecipitation analysis with antibodies recognizing indicated substance P (SP), KPI domain (KPI), C-terminal APP (M3) or Met-enkephalin (ME) epitopes; Lane M, molecular weight marker. B, Conditioned medium obtained from COS-1 cells expressing APP-REP 751 (lane 3), BAP<sub>E22Q</sub> (lanes 4, 6 and 8), BAP<sub>Δ11-28</sub> (lanes 5, 7 and 9) or control cells with (lane 2) or without (lane 1) transfection with vector DNA. C, Cell lysates obtained from COS-1 cells expressing APP-REP BAP<sub>E22Q</sub> (lanes 1, 4 and 7), BAP<sub>Δ11-28</sub> (lanes 2, 5 and 8) and control cells transfected with vector DNA (lanes 3, 6 and 9). D, Accumulation of secreted APP-REP 751 fragments in the conditioned medium obtained from COS-1 cells expressing APP-REP 751 (lanes 2 and 6), BAP<sub>E22Q</sub> (lanes 3 and 8), BAP<sub>Δ11-28</sub> (lanes 4 and 7), or control cells transfected with vector DNA (lanes 1 and 5), were pulsed with [<sup>35</sup>S]-methionine and chased for 45 (lanes 1-4) or 90 (lanes 5-8) minutes with cold methionine. E, Accumulation of secreted APP-REP fragments in the conditioned medium obtained from stable (Chinese hamster ovary cells; lanes 1-4) and transient (COS-1 cells; lanes 5 and 6) expression of APP-REP 751 (lanes 2 and 5), BAP<sub>Δ11-28</sub> (lanes 3 and 6), BAP<sub>E22Q</sub> (lane 4), or control cells transfected with vector DNA (lane 1).

Figure 6: Peptide mapping and sequencing of fragments secreted into the conditioned medium obtained from Chinese hamster ovary cells stably expressing APP-REP 751, BAP<sub>E22Q</sub> and BAP<sub>Δ11-28</sub>. A, Schematic representation depicting the APP-REP 751 and related derivative indicating the cleavage products and relevant carboxy-terminal fragments derived from treating the secreted fragments either with BNPS-Skatole (B) or cyanogen bromide. Downward- or upward-facing arrows represent BNPS-Skatole and cyanogen bromide cleavage sites, respectively. Amino acid lengths of relevant fragments for mapping or sequencing are given. B, BNPS-Skatole treatment of fragments secreted into the conditioned medium obtained from CHO cells stably expressing APP-REP 751 or BAP<sub>Δ11-28</sub>. Mixture of conditioned medium containing APP-REP and BAP<sub>Δ11-28</sub> (lane 1), or BAP<sub>Δ11-28</sub> (lane 2) and APP-REP 751 (lane 3) alone.

Figure 7: Nucleotide and amino acid sequence of the APP-REP 751 protein.

Figure 8: Nucleotide and amino acid sequence of APP 770 which also is available from the Genebank data base under accession number Y00264.

#### 45 DETAILED DESCRIPTION OF THE INVENTION

This invention provides a nucleic acid molecule encoding an amyloid precursor mutein, wherein the nucleic acid molecule comprises, from the 5' end to the 3' end, a nucleic acid sequence encoding a marker and a nucleic acid sequence encoding the amino terminus of APP up to but not including the sequences that encode BAP. These nucleic acid molecules may include, but are not limited to the nucleic acid molecules selected from the group consisting of pCLL983, pCLL935, pCLL934 and pCLL913.

This invention also provides a nucleic acid molecule encoding an amyloid precursor mutein, wherein the nucleic acid molecule comprises, from the 5' end to the 3' end a nucleic acid sequence encoding BAP and a nucleic acid sequence encoding a marker. These nucleic acid molecules may include, but are not limited to the nucleic acid molecules selected from the group consisting of pCLL947, pCLL914, pCLL937, pCLL949 and pCLL957.

Further provided by this invention is a nucleic acid molecule which comprises the nucleic acid molecules defined hereinabove to each other. Method of ligating are well known to those of skill in the art.

These nucleic acid molecules may include, but are not limited to the nucleic acid molecules selected from the group consisting of pCLL618, pCLL619, pCLL620, pCLL600, pCLL964, pCLL962, pCLL989, pCLL987, pCLL990, pCLL988, pCLL601, pCLL602, pCLL603, pCLL604, pCLL605, pCLL606 and pCLL607.

As used herein, the term "amyloid precursor mutein" is intended to encompass an amyloid precursor protein that is mutated, i.e., it is derived from a nucleic acid molecule which has changes in its primary structure as compared to wild-type amyloid precursor protein (APP). Wild-type APP exists in three isoforms, thus, the nucleic acid molecule is changed in its primary structure for each of the three isoforms of wild-type APP. As is known to those of skill in the art, a mutation may be a substitution, deletion, or insertion of at least one nucleotide along the primary structure of the molecule. The mutations which are encompassed by this invention are the result of saturation mutagenesis in the regions of APP which are susceptible to cleavage by endoproteolytic enzymes. These mutations include deletions of nucleic acids encoding particular amino acids, substitution of nucleic acid sequences encoding one amino acid for a different amino acid and addition of nucleic acid sequences encoding additional amino acids not present in the wild type APP sequence. The term "marker" encompasses any substance capable of being detected or allowing the nucleic acid or polypeptide of this invention to be detected. Examples of markers are detectable proteins, such as enzymes or enzyme substrates and epitopes not naturally occurring in wild-type APP that are capable of forming a complex with an antibody, e.g. a polyclonal or monoclonal antibody. In the preferred embodiment of this invention, the marker is an epitope that is capable of being detected by a commercially available antibody. In one embodiment, the marker is an epitope capable of being detected by a monoclonal antibody directed to the Substance P, the Met-enkephalin or the c-myc epitope. In the most preferred embodiment of this invention, the marker is the c-myc epitopic region.

The term "BAP region" is defined as the region of APP wherein endoproteolytic cleavage will yield the amino-terminus and the carboxy-terminus of the BAP which is deposited as plaques and cerebrovascular amyloid in Alzheimer's disease brain. The function of the "BAP region" is to give rise to BAP which may function as a neurotoxic and/or neurotrophic agent in the brain and as other functionalities ascribed to BAP. The "BAP region" may also be endoproteolytically cleaved by enzymes. Such enzymes may include, but are not limited to the enzymes multicatalytic prteinase, propyl-endopeptidase, Cathepsin-B, Cathepsin-D, Cathepsin-L, Cathepsin-G or secretase. Secretase cleaves between lysine-16 (K-16) and leucine-17 (L-17) where full length BAP comprises the amino acid sequence DAEFRHDSGYEVHHQKLVFFAEDVGSNK-GAIIGLMVGGVIA. Thus, for the purposes of this invention, the preferred embodiment is a cDNA which encodes an RNA which is translated into a protein which is the substrate for endoproteolytic activities which generate BAP.

In addition, for the purposes of this invention, the nucleic acid molecule may be DNA, cDNA or RNA. However, in the most preferred embodiment of this invention, the nucleic acid is a cDNA molecule.

This invention also encompasses each of the nucleic acid molecules described hereinabove inserted into a vector so that the nucleic acid molecule may be expressed, i.e., transcribed (when the molecule is DNA) and translated into a polypeptide in both procaryotic and eucaryotic expression systems. Suitable expression vectors useful for the practice of this invention include pSVL (Pharmacia), pRCRSV (Invitrogen), pBluescript SK<sup>+</sup> (Stratagene), pSL301 (Invitrogen), pUC19 (New England Biolabs). However, in the preferred embodiment of this invention, the vector pCDNA-1-neo is the expression vector for expression in eucaryotic cells. As is well known to those of skill in the art, the nucleic acid molecules of this invention may be operatively linked to a promoter of RNA transcription, as well as other regulatory sequences. As used herein, the term "operatively linked" means positioned in such a manner that the promoter will direct the transcription of RNA off of the nucleic acid molecule. An example of a promoter is the human cytomegalovirus promoter. The vectors of this invention preferably are capable of transcribing and/or translating nucleic acid in vitro or in vivo. The recombinant polypeptides produced from the expression of the nucleic acid molecules of this invention are also provided.

A host vector system for the production of the recombinant polypeptides described hereinabove and for expressing the nucleic acid molecules of the subject invention are provided. The host vector system comprises one of the vectors described hereinabove in a suitable host. For the purpose of the invention, a suitable host may include, but is not limited to a eucaryotic cell, e.g., a mammalian cell, a yeast cell or an insect cell for baculovirus expression. Suitable mammalian cells may comprise, but are not limited to Chinese hamster ovary cells (CHO cells), African green monkey kidney COS-1 cells, and ATCC HTB14 (American Type Tissue Culture). Most preferably, the cell lines CRL 1650 and CRL 1793 are used. Each of these are available from the American Type Culture Collection (ATCC) 12301 Parklawn Drive, Rockville, Maryland U.S.A. 20852. Suitable procaryotic cell may include, but are not limited to bacteria cells, HB101 (Invitrogen), MC1061/P3 (Invitrogen), CJ236 (Invitrogen) and JM109 (Invitrogen). Accordingly, the procaryotic or eucaryotic cell comprising the vector system of this invention is further provided by this

invention.

As is known to those of skill in the art, recombinant DNA technology involves insertion of specific DNA sequences into a DNA vehicle (vector) to form a recombinant DNA molecule which is capable of being replicated in a host cell. Generally, but not necessarily, the inserted DNA sequence is foreign to the recipient DNA vehicle, i.e., the inserted DNA sequence and DNA vector are derived from organisms which do not exchange genetic information in nature, or the inserted DNA sequence comprises information which may be wholly or partially artificial. Several general methods have been developed which enable construction of recombinant DNA molecules. For example, U.S. Patent No. 4,237,224 to Cohen and Boyer describes production of such recombinant plasmids using processes of cleavage of DNA with restriction enzymes and joining the DNA pieces by known method of ligation.

These recombinant plasmids are then introduced by means of transformation or transfection and replicated in unicellular cultures including prokaryotic organisms and eukaryotic organisms and eukaryotic cells grown in tissue culture. Because of the general applicability of the techniques described therein, U.S. Patent No. 4,237,224 is hereby incorporated by reference into the present specification. Another method for introducing recombinant DNA molecules into unicellular organisms is described by Collins and Hohn in U.S. Patent No. 4,304,863 which is also incorporated herein by reference. This method utilized a packaging, transduction system with bacteriophage vectors (cosmids).

Nucleic acid sequences may also be inserted into viruses, for example, a vaccinia virus or a baculovirus. Such recombinant viruses may be generated, for example, by transfection of plasmids into cells infected with virus, Chakrabarti et al, (1985) Mol. Cell Biol. 5:3402-3409.

Regardless of the method used for construction, the recombinant DNA molecule is preferable compatible with the host cell, i.e., capable of being replicated in the host cell either as part of the host chromosomes or as an extrachromosomal element. The recombinant DNA molecule or recombinant virus preferable has a marker function which allows the selection of the desired recombinant DNA molecule(s) or virus, e.g., baculovirus. In addition, if all of the proper replication, transcription and translation signals are correctly arranged on the recombinant DNA molecule, the foreign gene will be properly expressed in the transformed or transfected host cells.

Different genetic signals and processing events control gene expression at different levels. For instance, DNA transcription is one level, and messenger RNA (mRNA) translation is another. Transcription of DNA is dependent upon the presence of a promoter which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes RNA synthesis. The DNA sequences of eukaryotic promoters differ from those of prokaryotic promoters. Furthermore, eukaryotic promoters and accompanying genetic signals may not be recognized in or may not function in a prokaryotic system.

Similarly, translation of mRNA in prokaryotes depends upon the presence of the proper prokaryotic signals which differ from those of eukaryotes. Efficient translation of mRNA in prokaryotes requires a ribosome binding site called the Shine-Dalgarno (SD) sequence on the mRNA. For a review on maximizing gene expression, see Roberts and Lauer (1979) Methods in Enzymology 68:473.

Many other factors complicate the expression of foreign genes in prokaryotes even after the proper signals are inserted and appropriately positioned. One such factor is the presence of an active proteolytic system in *E. coli* and other bacteria. This protein-degrading system appears to destroy foreign proteins selectively. A tremendous utility, therefore, would be afforded by the development of a means to protect eukaryotic proteins expressed in bacteria from proteolytic degradation. One strategy is to construct hybrid genes in which the foreign sequence is ligated in phase (i.e., in the correct reading frame) with a prokaryotic structural gene.

Expression of this hybrid gene results in a recombinant protein product (a protein that is a hybrid of prokaryotic and foreign amino acid sequences).

Successful expression of a cloned gene requires efficient transcription of DNA, translation of the mRNA and in some instances post-translation modification of the protein. Expression vectors have been developed to increase protein production from the cloned gene. In expression vectors, the cloned gene is often placed next to a strong promoter which is controllable so that transcription can be turned on when necessary. Cells can be grown to a high density and then the promoter can be induced to increase the number of transcripts. These, if efficiently translated, will result in high yields of polypeptide. This is an especially valuable system if the foreign protein is deleterious to the host cell.

Several recombinant DNA expression systems are described below in the Experimental Procedures section for the purpose of illustration only, and these examples should not be construed to limit the scope of the present invention.

A method for producing a recombinant polypeptide described hereinabove, is also provided. This method comprises growing the host cell containing the nucleic acid of this invention and/or the host vector

system of this invention under suitable conditions, permitting production of the polypeptide and recovering the resulting recombinant polypeptide produced.

A method of detecting in a sample the presence of any of the recombinant polypeptides described hereinabove is further provided by this invention. In the preferred embodiment of this invention, the marker 5 is an epitope directed against an antibody, the epitope of which is not present in the wild-type polypeptide or APP derivative. This method comprises obtaining a sample suspected of containing the polypeptide and contacting the sample with an antibody directed to the marker. The contacting is done under suitable conditions to favor the formation of an antibody-epitope (i.e., antigen) complex, and detecting the presence of any complex so formed. The presence of complex being a positive indication that the recombinant 10 polypeptide is in the sample. In one embodiment of this invention, the antibody is a mouse antibody. In another embodiment of this invention, the antibody is a human antibody. In the most preferred embodiment, the mouse or human antibody is either a mouse or human monoclonal antibody.

The antibody is labeled with a detectable marker selected from the group consisting of radioisotopes, dyes, enzymes and biotin. For the purposes of this invention, suitable radioisotopes include, but are not 15 limited to,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{131}\text{I}$  and  $^{125}\text{I}$ .

Suitable samples for the practice of this invention include, but are not limited to conditioned media, cell lysates and cellular organelle fractions.

The method of this invention may utilize the recombinant polypeptide for the detection of drugs or compounds that inhibit or augment the activity of proteolytic enzymes which cleave APP to generate BAP 20 fragments. For the purposes of example only, a recombinant polypeptide which contains a Substance-P marker epitope on the amino-terminal side of BAP and a Met-enkephalin marker epitope on the carboxy-terminal side of BAP. Using commercially available RIA kits (Peninsula), one can measure the amount of amino-marker and carboxy-marker in any given sample. Since endoproteolytic activity is shown (see Figure 25 3) to allow the release of amino-terminal fragments of APP containing the amino-marker into the conditioned media while carboxy-terminal APP fragments containing the carboxy-marker remain associated with the cell, then RIA which measure the amount of amino-marker in the conditioned medium as a direct result of endoproteolytic cleavage activity between the marker epitopes preferable within the "BAP region". Using this RIA to the amino-marker, the effect of potential drugs designed to modify endoprotease activity can be tested comparing the level of amino-marker in untreated and endoprotease-inhibitor treated samples. If a 30 difference in non-treated and treated samples is found, then the position of the cleavage or lack of cleavage can be verified as with the procedures used in Figures 3 to 6. Thus, the qualitative and quantitative aspects of endoproteolytic activity and its inhibition on the recombinant APP murein is evaluated. The amino-marker also is an enzyme such as betagalactosidase which would be released int the conditioned media by the 35 action of an appropriate endoprotease. Cell free samples of conditioned media containing the liberated enzyme converts a chromogenic substrate into the appropriately colored product (Blue for X-gal and Yellow for ONPG) which is measured spectrophotometrically. Inhibitors of the appropriate endoprotease would inhibit the release of betagalactosidase enzyme into the conditioned medium resulting in less colored product being observed.

It is a purpose of this invention to develop a cleavable APP substrate system which represents target 40 sequences of BAP including normal flanking regions to provide recognition sequences for processing enzymes. The utilization of a common substrate for parallel strategies involving in vitro cleavage assays using cellular extracts in vivo processing assays in tissue culture or bacterial cells, or in conjunction with a selection system aimed at cloning BAP-cleaving proteases (or other relevant proteins) is preferred.

A second purpose of this invention is to develop an APP substrate which is non-cleavable by secretase 45 in order to better detect other putative abnormal processing events which are hypothesized to potentially either compete with secretase for limited substrate, or occur at much lower frequency than secretase and whose effects may be otherwise masked by the mass action of secretase.

Third, secretase-cleavable and -noncleavable APP substrates would provide probes with which to 50 investigate cellular posttranslational modifications to APP in an attempt to determine the potential influence on normal secretase and abnormal BAP "clipping" activities. These areas include, among others, the consideration of various known APP point mutations, contribution by different cell/tissue types (normal- or AD-specific), the Kunitz Protease Inhibitor domain present in APP-770 and -751 isoforms, APP phosphorylation and APP glycosylation.

Fourth, the ability to detect specific APP proteolytic events, either the normal secretase or the abnormal 55 BAP-generating activities, would enable the use of strategies which use phenotypic rescue as a marker for the cloning of potentially relevant and interesting proteases in tissue culture systems.

## Overview of the APP-REP Strategy

To study secretase and BAP-generating pathways, portions of APP cDNA clones are used to engineer a panel of APP-REPorter (APP-REP) plasmids to express "marked" proteins representing each of the APP isoforms (and other APP/BAP sequence alterations; see below) in cultured cells. The system utilizes the marker Substance-P (SP) and Met-Enkephalin (ME) which are strategically placed, respectively, on amino- and carboxy-terminal sides of BAP. Proteolytic cleavage of APP-REP target substrate is determined by the electrophoretic sizing of resulting proteolytic fragments and immunological detection of APP-specific and SP and ME reporter epitopes. Deletion of a large central portion of APP sequence also makes APP-REP readily distinguishable from the endogenous APP isoforms based on size. Moreover, the resolution of detecting proteolytic cleavage at different positions within the APP-REP substrate protein is enhanced by working with shorter target substrates. Approximate location of cleavage is determined initially by fragment sizing and epitope mapping; the exact cleavage site is later determined by peptide mapping of affinity/HPLC purified fragments and sequencing of peptide ends.

Plasmids also are derived from these constructs for developing similar strategies to express APP-REP protein in cell free reticulocyte transcription-translation and bacterial systems. Mutation of APP-REP secretase/BAPase cleavage site (by sequence substitution, deletion or FAD mutations) can reveal putative proteolytic activities associated with BAP formation including amino- and carboxy-BAPase activities which are predicted to result in altered product fragments lengths.

20

## FIRST SERIES OF EXPERIMENTS

### Bacterial Strains and Transformation

Transformation of commercially available frozen competent bacteria, maintenance and selection of transformants is according to the manufacturer. Strains HB101, DH5a or JM109 (Gibco-BRL) are used for the construction of APP-REP in pSK(+) (Stratagene, La Jolla, CA) and pSL 301 (Invitrogen, San Diego, CA). APP-REP is subsequently subcloned into the eucaryotic expression vector pcDNA-1-neo and amplified in MC1061/P3 (Invitrogen, San Diego, CA).

30

### Plasmid Construction

A cassette approach is used to independently construct portions of the APP-REP plasmid (Figure 2). The N-terminal partial includes APP sequences through the Substance P (SP) epitope, while the carboxy-terminal (C-terminal) partial includes BAP (or sequence variations of BAP) through the Metenkephalin (ME) epitope (Figure 1). Plasmid encoding the N-terminal cassette (pCLL935) is constructed by ligating the EcoRI-Xhol fragment derived from APP-751 cDNA to a short synthetic Xhol-HindIII fragment encoding Substance P (amino acid 1-11). This product is then ligated into the EcoRI and HindIII sites of pSK(+). Plasmid encoding the carboxy-terminal (C-terminal) cassette (pCLL947) is constructed by cloning into the HindIII-BamHI sites of pSL301 a fragment containing BAP sequences which is amplified by polymerase chain reaction. The fragment features a novel 5'-HindIII site beginning at lysine 638 of APP-751, native BAP through APP C-terminal sequences, and a C-terminal fusion including the Metenkephalin epitope followed by a stop translation codon and a BamHI site. The resulting pSL301 HindIII-Sall fragment (including the HindIII-BamHI coding region plus BamHI-Sall polylinker sequences) is then isolated and ligated to the N-terminal cassette by subcloning into the HindIII-Sall sites of the SK(+)-based, CMV promoter driven, eukaryotic expression vector pcDNA-1-neo (pCLL601), whose polylinker is modified to accommodate the APP-REP fragment (pCLL602). Polylinker modification involves the substitution of the HindIII-XbaI fragment with a synthetic one which restores HindIII, destroys XbaI and introduces novel BamHI-XbaI-Xho-Sall sites.

50

### Tissue Culture Lines

All cells are obtained from American Type Culture Collection and maintained according to their recommendation. They include SV40-transformed African Green monkey kidney COS-1 cells (CRL 1650) for transient expression and Chinese hamster ovary CHO-1C6 (CRL 1973) for stable expression systems.

55

## Transfection Procedure

- Cells are seeded at a density of 2-3 X 10<sup>5</sup>/100 mm dish and transfected using Lipofectin (Gibco-BRL, Grand Island, NY) when ~75% confluent. Plasmid DNA (0.5-4 mg) is diluted in 450 ml of Opti-MEM (Gibco-BRL, Grand Island, NY), mixed with 450 ml containing 75-100 ml Lipofectin and the mixture incubated at room temperature for 20-30 minutes. Addition of DNA-Lipofectin mixture to cells, recovery phase and G418 selection (Gibco-BRL), when applicable, are according to the manufacturer's protocol. Cells and conditioned medium are harvested at 48-72 hours following transfection for assay of APP-REP expression.

## 10 Antisera

APP-specific antisera:anti-N-terminal APP, mouse monoclonal 22C11 (Boehringer-Mannheim Biochemicals, Indianapolis, IN) raised against a recombinant fusion protein expressing APP-695 (epitope mapped to aa 60-100); anti-KPI rabbit polyclonal, raised against recombinant protein encoded by the Hinfl fragment derived from APP-770; and anti-APP C-terminal rabbit polyclonal M3, raised against synthetic APP peptides corresponding to APP-770 amino acid residues 649-671. Reporter-specific antisera:anti-substance P, rabbit polyclonal, purchased from Peninsula, Belmont, CA; and anti-Met-enkephalin, rabbit polyclonal, purchased from Cambridge, Wilmington, DE.

## 20 Preparation of Radiolabeled APP-REP and Extraction from Conditioned Medium and Cell Lysates

APP-REP proteins transiently expressed in exponentially growing adherent cells (~4 x 10<sup>6</sup>) are radiolabeled by metabolic incorporation of [<sup>35</sup>S]-methionine as follows. Cell monolayers are washed twice with prelabeling medium (methionine-free D-HEM supplemented with glutamine, sodium pyruvate, antibiotics and 1% dialyzed fetal bovine serum (Gibco-BRL) and incubated for 15 minutes to 4 hours in prelabeling medium containing 150-450 uCi[<sup>35</sup>S]-methionine (Amersham, 800Ci/mmol). If chased with cold methionine, the medium is removed following the pulse, the monolayer is washed with prelabeling medium and replaced with 3 ml of the same containing 1 mM cold methionine.

The conditioned medium is recovered following radiolabeling by aspiration from plates and cell debris removed by centrifugation for 10 minutes at 4°C (~300xg). Conditioned medium is immediately supplemented with protease inhibitors (pepstatin A, 50 ug/ml; leupeptin, 50 ug/ml; aprotinin, 10 ug/ml; EDTA, 5 mM; PMSF, 0.25 mM) and immunoprecipitation buffer (IPB; Sisodia *et al.*, 1990) for protein analysis. Briefly, 3 ml of CM is supplemented with 0.75 ml 5X IPB (250 mM Tris, pH 6.8; 750 mM NaCl; 25 mM EDTA; 2.5% Nonidet P40; 2.5% sodium deoxycholate) and incubated for 20 minutes at 4°C prior to use.

35 Lysates are prepared by washing the labeled cell monolayer twice with 5 ml pre-labeling medium and directly extracting cells in plates at 4° C with 3.75 ml 1X IPB (including protease inhibitors). Cells are scraped into the buffer, incubated for 20 minutes at 4°C and lysates clarified of cellular debris by centrifugation for 20 minutes at 10,000xg.

For radioiodination of cell surface proteins, monolayers are chilled on ice, washed 3 times with 5 ml ice cold PBS and labeled at room temperature for 10 minutes following the addition of: 5 ml PBS containing 0.2 mCi Iodine-125 (NEZ-033A, New England Nuclear), 0.25 ml lactoperoxidase (1 mg/ml distilled water, Sigma), 10 ul of hydrogen peroxide solution (freshly prepared by diluting 10 ml of 30% stock in 10 ml of PBS) added at 0, 3, 6, and 9 minutes of iodination. At 10 minutes, the supernatant is removed and cells gently washed with 10 ml of ice cold PBS (containing 10 mM NaI). Four ml of PBS is added, and CM and cell lysates are prepared as above.

## Immunoprecipitation Analysis

Aliquots of radiolabeled lysate or conditioned medium representing 4-8x10<sup>5</sup> cells are thawed on ice, supplemented with protease inhibitors (see above), boiled for 3 minutes in 0.35% SDS and chilled on ice. Samples are preincubated on a shaker for 1.5 hours at 4°C with 2-10 ul 2X of preimmune (or normal rabbit) serum and 2 mg Protein A-Sepharose (Sigma; prepared in 1X IPB), and insoluble immune removed by centrifugation. APP-or reporter epitope-specific antisera (0.1-10 ul) and 2 mg Protein A-Sepharose were similarly added and incubated overnight. Specific immune complexes were precipitated, washed 4 times with 0.25 ml 1 X IPB (with protease inhibitors), extracted with 20 ul Laemmli sample buffer (Laemmli (1970) Nature 227:680-685), boiled for 3 minutes and fractionated by electrophoresis on SDS-polyacrylamide-tris-glycine (Bio-Rad Laboratories, Richmond, VA) or SDS-polyacrylamide-tris-tricine Daiichi (Integrated Separation Systems, Natick, MA) gels. Gels are then treated with Enlightening Autoradiographic Enhancer (New

England Nuclear, NEF-974) and dried in vacuo with heat and exposed to Kodak X-AR film at -70°C.

#### Western (Immunoblot) Analysis

- 5 Lysate or 10X concentrated conditioned medium (Centricon 30 microconcentrator; Amicon, Beverly, MA) representing 4-8x10<sup>5</sup> cells are supplemented with an equal volume of 2X Laemmli sample buffer, boiled for 2 minutes, fractionated by electrophoresis on SDS-polyacrylamide-tris-glycine (Bio-Rad, XX) or SDS-polyacrylamide-tris-tricine Daiichi (Integrated Separation Systems, Natick, MA) gels and transblotted (Semi-Phor, Hoefer Instruments, San Francisco, CA) to Immobilon-P membrane (Millipore, Bedford, MA).
- 10 Membranes are pre-blocked in 10 ml 5% non-fat dry milk/PBST (PBS with 0.02% Tween-20) for 45 minutes at room temperature prior to overnight incubation at 4°C with primary antisera (in fresh pre-blocking solution). Blots are then washed, incubated with secondary antibody, washed and developed for horseradish peroxidase activity as described (ECL Luminol Kit; Amersham, Arlington Heights, IL).

#### 15 Peptide Mapping and Determination of the Site of Proteolytic Cleavage by Peptide Sequencing

The secretase clip site is determined essentially as described (Wang et al., (1991) J. Biol. Chem. 266:16960-16964). Approximately 1X10<sup>6</sup> CHO cells stably expressing APP-REP are seeded in each 150 mm dish containing DMEM (complete with 200 ug/ml G418) and incubated for 36 hours. Cells are washed, preincubated for 6 hours in serum-free medium [MCDB 302 supplemented with antibiotics, L-glutamine (292 mg/l) and proline 12 mg/l (Sigma) to remove serum components, washed, and incubated for another 72 hours in fresh serum-free media.

Serum-free conditioned medium was pooled and cell debris is removed by centrifugation (10 minutes at 300xg, then 30 minutes at 100,000xg) and concentrated by acetone precipitation and fractionated by FPLC. Conditioned medium concentrate is loaded on an anion exchange column (Mono Q; source) and protein is eluted in 20 mM Tris (pH 7.4) over a 0-1M NaCl gradient. Fractions containing secreted APP are identified by immunoblotting (monoclonal antibody 22C11) and relevant samples pooled, desalted (NP-5 column; Pharmacia, Piscataway, NJ) and concentrated. Proteins are then denatured, treated with cyanogen bromide (in 10% trifluoroacetic acid) and peptides separated by high performance liquid chromatography (Vydac C<sub>18</sub> reverse-phase) attached to a FAB-MS unit. Relevant peaks derived from APP-REP 751 and APP-REP BAP<sub>11-28</sub> are identified by locating those peaks uncommon to both proteins. The C-terminal peptides derived from APP-REP BAP<sub>11-28</sub> (predicted 14 amino acid) and APP-REP 751 (predicted 17 amino acid) are sequenced (MilliGen solid phase peptide sequencer; Millipore, Burlington, MA).

#### 35 EXPERIMENTAL RESULTS

##### Characterization of APP-REP Expression by Epitope Mapping

The APP-REP strategy (Figure 1) is system for the expression of marked APP proteins in tissue culture cells in order to characterize the proteolytic cleavage events. The deletion of 276 amino acid portion distinguishes the construct of this invention from endogenously expressed APP on the basis of size, and is predicted to increase the resolution of APP-REP fragments resulting from the proteolytic cleavage by secretase or other amyloidogenic, BAP-generating cleavage events. Substance P and Met-enkephalin marker epitopes strategically placed on either side of BAP enable the immunological detection of N- and C-terminal fragments, respectively, which result from proteolytic cleavage of APP-REP substrate.

APP-REP protein transiently expressed in COS-1 cells has been radiolabeled by metabolic incorporation of [<sup>35</sup>S]-methionine in a 60 minute pulse, immunoprecipitated with antisera, and size fractionated by gel electrophoresis as demonstrated in Figure 3. Immunoprecipitation with a panel of APP- and APP-REP-specific antisera which recognize epitopes mapping at various positions along APP-REP, reveals the presence of 2 proteins of ~63 kDa in cell lysates (including cytoplasmic and membrane associated proteins) as shown in Figure 3. The specific detection by antisera directed against the KPI domain, the carboxy-terminus of APP (M3, Figure 3A) and Met-enkephalin, as well as by the N-terminal 22C11 monoclonal in Western blot analysis (data not shown), suggests that both bands represent the full-length APP-REP protein. Although the 492 amino acid APP-REP is predicted to display a mobility of ~49-54 kDa, the larger 63 and 55 76 kDa proteins are expected based on previous observations attributing the aberrant migration properties of APP, putatively to post-translational modification like tyrosine-sulfation, glycosylation and phosphorylation (Dyrks et al., (1988) EHSO J. 7:949-957; Weidemann et al., (1989) Cell 57:115-126.

Analysis of the conditioned medium (CM) collected from those same cells above indicates that an N-terminal fragment of APP-REP is released into the CM. Figure 3B reveals a shorter ~67 kDa fragment immunoprecipitable from CM with KPI and SP antisera (and the 22C11 monoclonal by Western analysis), but not with several C-terminal APP or ME antisera. These data are consistent with the observations (Selkoe et al., (1988) P.N.A.S. 86:6338-6342; Palmert et al., (1989) a) P.N.A.S. U.S.A. 85:7341-7345), b) indicating that APP is a substrate for the proteolytic cleavage resulting in the secretion of an N-terminal fragment into CM, and a short membrane associated C-terminal fragment.

Pulse-Chase Analysis Reveals the Precursor/Product Relationship Between Cell Associated and Secreted Derivatives of APP-REP

To show that APP-REP undergoes post-translational modification accounting for the 2 cell associated proteins, and that the N-terminal APP-REP fragment released into CM is derived from one of these precursors, radiolabeled APP-REP is with a short 15 minute pulse and collected both cell lysates and CM at various chase intervals as shown in Figure 4. Immunoprecipitation analysis reveals that APP-REP initially migrates at ~63 kDa and is rapidly "chased" up to ~76 kDa with conversion rate of less than 10-15 minutes (Figure 4A; also see Figure 5C for quantitative analysis), an observation which is consistent with the notion that APP-REP, like APP, is substrate for posttranslational modifications.

The ~76 kDa APP-REP band (cell lysate) rapidly disappears ( $t^{1/2} \sim 20$  minutes) (Figure 4A and 5C), followed by the appearance of a shorter ~67 kDa band in the CM (Figure 4B and 5C). The released ~67 kDa fragment accumulates rapidly and is relatively long lived ( $t^{1/2} > 8$  hours). The temporal pattern of intracellular APP-REP depletion, accumulation of a shorter ~67 kDa protein in CM, and the recognition of this protein only by antisera raised against N-terminal epitopes, is consistent with proteolytic cleavage of APP-REP which is similar to the normal, non-amyloidogenic, "secretase" activity which results in the release of an N-terminal APP fragment (Sisodia et al., (1990) Science 248:492-495).

Expression of APP-REP Derivatives Containing Altered BAP Sequences Does Not Prevent Proteolytic Cleavage

In an attempt to engineer non-cleavable substrates for secretase, APP-REP proteins are expressed (Figure 5A) either lacking the secretase "cleavage/recognition site" putatively encompassed by aa residues BAP 11-28 (BAP<sub>Δ11-28</sub>pCLL604), or representing the BAP point mutation found in patients with HCHWA-D (BAP<sub>E22Q</sub>pCLL603). The construct representing the BAPE22Q mutation results in secretion of an N-terminal fragment indistinguishable from the APP-REP protein (Figure 5C). Deletion of extracellular, juxtamembranous 18 aa (BAP<sub>Δ</sub> 11-28), however, still results in the secretion of an N-terminal APP-REP fragment into the CM (Figure 5B). A slightly faster migration of fragment derived from the deletion construct pCLL604 in comparison to that of wild-type pCLL602, is consistent with the 18 aa deletion and a corresponding loss of ~2 kDa (Figure 5C). Pulse-chase analyses (Figure 5D) indicate that expression of full-length precursor by each construct, proteolytic cleavable and the release of fragment into CM is both qualitatively and quantitatively similar to that of the wild-type APP-REP sequence. Chinese hamster ovary (CHO) cells stably expressing APP-REP display results similar to that of transiently expressing COS-1 cells (Figure 5E). Collectively, these data suggest that the cleavage in each case may be the result of similar biochemical events despite the difference in juxtamembranous sequences (Figure 5A).

Full-Length APP-REP Proteins Are Associated with Plasma Membrane Prior to Cleavage

In preliminary experiments, detection of the amino-terminal APP-REP fragment in CM and not in cell lysates, suggests that the putative secretase activity might be plasma membrane-associated. One prediction of this notion is that an N-terminal portion of APP-REP might be (partially) localized to the extracellular environment prior to cleavage. In order to test this hypothesis, CHO cells stably expressing APP-REP (pCLL602) are subjected to lactoperoxidase-catalyzed iodination to radiolabel only extracellular proteins associated with the cell surface, and CM and cell lysates were analyzed immediately following iodination or after a 10 minute incubation. Presence of the ~76 kDa APP-REP band in cell lysate should indicate that at least a portion of full-length APP-REP is poised extracellularly in association with cell membrane. Detection of both, a reduced fraction of the ~76 kDa band in the cell lysate and a corresponding increased fraction of ~67 kDa fragment in CM following the "release" incubation, would suggest that the extracellular portion of APP-REP is cleaved.

### Peptide Sequencing to Determine the site of Proteolysis

Fragment secreted into serum-free media derived from CHO cells stably expressing APP-REP with wild-type or BAP 11-28 sequences has been analyzed to determine the actual site of proteolytic cleavage as shown in Figure 6. Peptide mapping by tryptophan-specific cleavage with BNPS-skatole is used to roughly determine the approximate position of cleavage in each molecule. Western blot analysis using SP antisera following BNPS-skatole treatment (Figure 6B) reveals fragments whose lengths of ~10.5 and ~9.5 kDa, corresponding to wild type and BAP<sub>11-28</sub> respectively, confirming that cleavage occurs in the C-terminal portion of the PN-2-like protein as expected (Figure 6A). To determine the actual position of cleavage, secreted fragment is partially purified, treated with cyanogen bromide and relevant C-terminal peptides derived from APP-REP wild type.

### DISCUSSION

The examined the expression of a truncated form of APP-751, namely APP-REP 751 (pCLL602) is examined and its normal cleavage by secretase. Comparison of the nontransfected cells and those transfected with APP-REP 751, in both COS-1 transient and CHO stable expression systems, show the production of shorter secreted protein derived from APP-REP. Furthermore, upon a prolonged exposure of the fluorogram only one band is observed in condition medium. Epitope mapping with antibodies to N- and C-terminal domains of APP-REP and amino acid sequencing suggest post-translational cleavage at a site similar to that reported for intact APP protein and other truncated APP constructs similar to that reported for intact APP protein and other truncated APP constructs. Pulse-chase experiments reveal post-translational modifications, believed to be similar to those described for the intact APP protein, in which a single ~63 kDa product is chased up to ~76 kDa in the first 30 minutes. Appearance of the ~76 kDa cell membrane associated protein precedes the release of a ~67 kDa product into the CM. The released form, which is not observed in the cell lysate fraction, steadily accumulates in the conditioned medium well after the ~76 kDa band has begun to disappear suggesting a precursor-product relationship. These data indicate that the APP-REP protein is a good representation of the naturally occurring APP with respect to post-translational synthesis, processing, and stability in a tissue culture system.

Epitope mapping of APP-REP 751 mutants suggest that BAP<sub>E22Q</sub>, as well as the BAP<sub>Δ11-28</sub> deletion constructs, are initially expressed as larger proteins of predicted lengths which subsequently are cleaved to release N-terminal fragments into the CM. The pulse-chase experiments indicate the cell-associated and secreted forms accumulate with similar kinetics.

35

40

45

50

55

TABLE 1  
Construction of APP-REP Partials

5      A.    pSK(+) Amino-Terminal Constructs:  
 Cloning of APP Isoform and Reporter  
Epitope (EcoRI-HindIII Fragments)

10	Plasmid	APP Isoform	Reporter Epitope
	Name	(EcoRI-XhoI Fragment)	(XhoI-
		HindIII Fragment)	
15	pCLL983	APP-695	Substance P*
	pCLL935	APP-751	Substance P
	pCLL934	APP-770**	Substance P
	pCLL913	APP-770#	Substance P

20      Notes:

\*      Substance P is a peptide containing 11 residues with  
 the amino acid sequence of RPKPQQFFGLM.

25      \*\*     5' untranslated sequences derived from the shorter APP-  
 770 cDNA form.

#      5' untranslated sequences derived from the longer APP-  
 751 cDNA form.

30      B.    pSL301 Carboxy-Terminal Constructs: Cloning  
 of BAP-Encoding APP Reporter Epitope Fusions  
(HindIII-BamHI/SalI Fragment)

40	Plasmid	Met-Enkephalin (ME)	
	Name	Fusion at end of:	Name of Variation
	pCLL947	Full-Length APP	APP-BAP-APP-ME
	pCLL914	Transmembrane Domain	APP-BAP-TM-ME
	pCLL937	BAP	APP-BAP-ME

45

50

55

TABLE 1  
Construction of APP-REP Partials  
(Continued)

5 C. pSL301 Carboxy-Terminal Full-Length APP-ME  
Constructs: Introduction of Mutations in BAP  
(HindIII-BamHI/SalI Fragment)

10	Plasmid	Met-Enkephalin	
	<u>Name</u>	<u>Fusion at End of:</u>	<u>Name of Variation</u>
	pCLL949	E to Q substitution at BAP aa#22	BAP22EQ
15		G to A substitution at BTaa#10, deletion of BAP AA#11-28 and creation of NdeI site	BAP-vaall-28
20			

25

30

35

40

45

50

55

TABLE 2  
**Assembly of APP-REP full-Length Constructs  
 Containing Substance P and Met-Enkephalin  
 Reporter Epitopes and BAP or a Variation of BAP**

Plasmid Name	Construct	Plasmid (N-Terminus)	Fragment (C-Terminus)	Restriction
pCLL618	APP-REP-695	pCLL983		pCLL947
pCLL964	APP-REP-751	pCLL935		pCLL947
pCLL962	APP-REP-770	pCLL934		pCLL947
pCLL619	APP-REP-695/BAP <sub>E</sub> to Q	pCLL983		pCLL949
pCLL989	APP-REP-751/BAP <sub>E</sub> to Q	pCLL935		pCLL949
pCLL987	APP-REP-770/BAP <sub>E</sub> to Q	pCLL934		pCLL949
pCLL620	APP-REP-695/BAP <sub>a11-28</sub>	pCLL983		pCLL957
pCLL990	APP-REP-751/BAP <sub>a11-28</sub>	pCLL935		pCLL957
pCLL988	APP-REP-770/BAP <sub>a11-28</sub>	pCLL934		pCLL957

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

TABLE 3  
Subcloning of APP-REP Full-Length Constructs  
and Human Growth Hormone (hGH) into pcDNA-1-Neo[XS1]

<u>Plasmid Name</u>	<u>Construct Name (in pcDNA-1-neo)</u>	<u>Source of Insert</u>
pCLL600	pcDNA-1-neo-hGH	pOGH*
pCLL601	pcDNA-1-neo[XS]	Synthetic Fragment**
pCLL602	APP-REP-751	pCLL964
pCLL603 #	APP-REP-751/BAP <sub>E</sub> to Q	pCLL989
pCLL604 #	APP-REP-751/BAPaa11-28	pCLL990
pCLL605	APP-REP-770	pCLL962
pCLL606	APP-REP-770/BAP <sub>E</sub> to Q	pCLL987
pCLL607	APP-REP-770/BAPaa11-28	pCLL988

Notes:

- \* The HindIII-EcoRI (blunt-ended) fragment encoding hGH sequences of pOGH (Nichols Diagnostics) was subcloned into the HindIII-EcoRI (blunt-ended) sites of pcDNA-1-neo.
- \*\* The HindIII-XbaI fragment of the pcDNA-1-neo polylinker was replaced with a synthetic fragment which destroyed the original XbaI site and introduced several unique sites (HindIII-BamHI-XbaI-XbaI-SalI).
- # Also created by an alternative strategy using the same pSK(+) plasmids.

5

10

15

20

25

30

35

40

45

50

55

TABLE 4  
 "Secretase-Minus" APP-REP Constructs  
Engineered by Oligonucleotide-Directed Mutagenesis

<u>Plasmid Name</u>	<u>Mutation Identity</u>	<u>Mutated BAP Sequence Compared to Wild Type</u>				<u>Percent** Secretion</u>
pCLL602 BAP*	BAP*	CAT	CAA	AAA		100
		H	Q	K	L V F F	
pCLL608 BAP-16KE	BAP-16KE	CAT	CAA	<u>GAG</u>	TTG GTG TTC TTT	0
		H	Q	E	L V F F	
pCLL609 BAP-16KV	BAP-16KV	CAT	CAA	<u>GTG</u>	TTG GTG TTC TTT	10-20
		H	Q	V	L V F F	
pCLL610 BAP-19FP	BAP-19FP	CAT	CAA	AAA	TTG GTG <u>CCG</u> TTT	10-20
		H	Q	K	L V P F	

## Notes:

\* Wild-type BAP

\* % secretion relative to wild type BAP sequence as determined by Sisodia.

TABLE 5  
APP-REP Constructs Modeling APP Mutations  
Associated with Diseases Involving BAP Deposition

5

APP "717" MUTATIONS

10

// APP Transmembrane Domain //										
<u>// [BAP]</u>										
		711	712	713	714	715	716	717	718	719
		[40	41	42]						
15	pCLL602 APP*	GTC	ATA	GCG	ACA	GTG	ATC	GTC	ATC	ACC
		V	I	A	T	V	I	V	I	T
	pCLL611 717VI**	GTC	ATA	GCG	ACA	GTG	ATC	<u>ATC</u>	ATC	ACC
		V	I	A	T	V	I	I	I	T
20	pCLL612 717VG@	GTC	ATA	GCG	ACA	GTG	ATC	<u>GTC</u>	ATC	ACC
		V	I	A	T	V	I	G	I	T
	pCLL613 717VF\$	GTC	ATA	GCG	ACA	GTG	ATC	<u>TTC</u>	ATC	ACC
25		V	I	A	T	V	I	F	I	T'

30

TABLE 5 (continued)

35

DUTCH DISEASE	: V (secretase clip)										
	686	687	:	688	689	690	691	692	693	694	
	[15	16	:	17	18	19	20	21	22	23]	
40	pCLL602 BAP*	CAA	AAA	:	TTG	GTG	TTC	TTT	GCG	GAA	GAT
		Q	K	:	L	V	F	F	A	E	D
	pCLL603 BAP-22EQ#	CAA	AAA	:	TTG	GTG	TTC	TTT	GCA	<u>CAA</u>	GAT
45	pCLL606#	Q	K	:	L	V	F	F	A	Q	D

## Notes:

- 50 # APP-REP-751 and -770 derived BAP-22EQ constructs.
- \*\* Goate et al. (1991) Nature, 349:704-706; Yoshioka et al. (1991) BBRC 178:1141-1146; Naruse et al. (1991) Lancet 337:978-979.
- 55 @ Chartier-Harlin et al. (1991) Nature 353:844-846.
- \$ Murrell et al. (1991) Science 254:97-99.

(2) INFORMATION FOR SEQ ID NO:27:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: protein

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Gln Lys Leu Val Phe Phe Ala Gln Asp  
1 5

15

20

25

30

35

40

45

50

55

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

5

## (i) APPLICANT:

- (A) NAME: American Cyanamid Company
- (B) STREET: 1937 West Main Street
- (C) CITY: Stamford
- (D) STATE: Connecticut
- (E) COUNTRY: U.S.A
- (F) POSTAL CODE (ZIP): 06904-0060

10

(ii) TITLE OF INVENTION: Novel Amyloid Precursor Proteins and Methods  
of Using Same

15

(iii) NUMBER OF SEQUENCES: 27

20

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (2) INFORMATION FOR SEQ ID NO:1:

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1721 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

40

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 196..1671

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAGCTTGGGG ATCCGCTCTA GAACTAGTGG ATCCCCCGGG CTGCAGGAAT TCGGGGGGGG

60

45

50

55

	CAGCGGTAGG CGAGAGCACG CGGAGGAGCG TGCAGGGGGC CCCGGGAGAC GGCGGCCGTG	120
	GCGGCGCGGG CAGAGCAAGG ACGCGGCGGA TCCCCTCGC ACAGCAGCGC ACTCGGTGCC	180
5	CCGCGCAGGG TCGCG ATG CTG CCC GGT TTG GCA CTG CTC CTG CTG GCC GCC Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala 1 5 10	231
	TGG ACG TCT CGG GCG CTG GAG GTA CCC ACT GAT GGT AAT GCT GGC CTG Trp Thr Ser Arg Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu 15 20 25	279
10	CTG GCT GAA CCC CAG ATT GCC ATG TTC TGT GGC AGA CTG AAC ATG CAC Leu Ala Glu Pro Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His 30 35 40	327
15	ATG AAT GTC CAG AAT GGG AAG TGG GAT TCA GAT CCA TCA GGG ACC AAA Met Asn Val Gln Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys 45 50 55 60	375
	ACC TGC ATT GAT ACC AAG GAA GGC ATC CTG CAG TAT TGC CAA GAA GTC Thr Cys Ile Asp Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val 65 70 75	423
20	TAC CCT GAA CTG CAG ATC ACC AAT GTG GTA GAA GCC AAC CAA CCA GTG Tyr Pro Glu Leu Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val 80 85 90	471
25	ACC ATC CAG AAC TGG TGC AAG CGG GGC CGC AAG CAG TGC AAG ACC CAT Thr Ile Gln Asn Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His 95 100 105	519
	CCC CAC TTT GTG ATT CCC TAC CGC TGC TTA GTT GGT GAG TTT GTA AGT Pro His Phe Val Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser 110 115 120	567
30	GAT GCC CTT CTC GTT CCT GAC AAG TGC AAA TTC TTA CAC CAG GAG AGG Asp Ala Leu Leu Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg 125 130 135 140	615
35	ATG GAT GTT TGC GAA ACT CAT CTT CAC TGG CAC ACC GTC GCC AAA GAG Met Asp Val Cys Glu Thr His Leu His Trp His Thr Val Ala Lys Glu 145 150 155	663
	ACA TGC AGT GAG AAG AGT ACC AAC TTG CAT GAC TAC GGC ATG TTG CTG Thr Cys Ser Glu Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu 160 165 170	711
40	CCC TGC GGA ATT GAC AAG TTC CGA GGG GTA GAG TTT GTG TGT TGC CCA Pro Cys Gly Ile Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro 175 180 185	759
45		

50

55

	CTG GCT GAA GAA AGT GAC AAT GTG GAT TCT GCT GAT GCG GAG GAG GAT Leu Ala Glu Glu Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp 190 195 200	807
5	GAC TCG GAT GTC TGG TGG GGC GGA GCA GAC ACA GAC TAT GCA GAT GGG Asp Ser Asp Val Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly 205 210 215 220	855
10	AGT GAA GAG AAA GTA GTA GAA GTA GCA GAG GAG GAA GAA GTG GCT GAG Ser Glu Glu Lys Val Val Glu Val Ala Glu Glu Glu Val Ala Glu 225 230 235	903
15	GTC GAA GAA GAA GAA GCC GAT GAT GAC GAG GAC GAT GAG GAT GGT GAT Val Glu Glu Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp 240 245 250	951
20	GAG GTA GAG GAA GAG GCT GAG GAA CCC TAC GAA GAA GCC AGA GAG AGA Glu Val Glu Glu Glu Ala Glu Glu Pro Tyr Glu Glu Ala Arg Glu Arg 255 260 265	999
25	ACC ACC AGC ATT GCC ACC ACC ACC ACC ACC ACA GAG TCT GTG GAA Thr Thr Ser Ile Ala Thr Thr Thr Thr Thr Glu Ser Val Glu 270 275 280	1047
30	GAG GTG GTT CGA GAG GTG TGC TCT GAA CAA GCC GAG ACG GGG CCG TGC Glu Val Val Arg Glu Val Cys Ser Glu Glu Ala Glu Thr Gly Pro Cys 285 290 295 300	1095
35	CGA GCA ATG ATC TCC CGC TGG TAC TTT GAT GTG ACT GAA GGG AAG TGT Arg Ala Met Ile Ser Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys 305 310 315	1143
40	GCC CCA TTC TTT TAC GGC GGA TGT GGC GGC AAC CGG AAC AAC TTT GAC Ala Pro Phe Phe Tyr Gly Gly Cys Gly Asn Arg Asn Asn Phe Asp 320 325 330	1191
45	AGA GAA GAG TAC TGC ATG GCC GTG TGT GGG AGC GCC ATT CCT ACA ACA Arg Glu Glu Tyr Cys Met Ala Val Cys Gly Ser Ala Ile Pro Thr Thr 335 340 345	1239
50	GCA GCC AGT ACC CCT GAT GCC GTT GAC AAG TAT CTC GAG CGG CCC AAG Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu Glu Arg Pro Lys 350 355 360	1287
55	CCC CAG CAG TTC TTT GGC CTG ATG GGA AGC TTG ACA AAT ATC AAG ACG Pro Gln Gln Phe Phe Gly Leu Met Gly Ser Leu Thr Asn Ile Lys Thr 365 370 375 380	1335
60	GAG GAG ATC TCT GAA GTG AAG ATG GAT GCA GAA TTC CGA CAT GAC TCA Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser 385 390 395	1383

45

50

55

GGA TAT GAA GTT CAT CAT CAA AAA TTG GTG TTC TTT GCA GAA GAT GTG Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val 400 405 410	1431	
5 GGT TCA AAC AAA GGT GCA ATC ATT GGA CTC ATG GTG GGC GGT GTT GTC Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val 415 420 425	1479	
10 ATA GCG ACA GTG ATC GTC ATC ACC TTG GTG ATG CTG AAG AAG AAA CAT Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys His 430 435 440	1527	
TAC ACA TCC ATT CAT CAT GGT GTG GTG GAG GTT GAC GCC GCT GTC ACC Tyr Thr Ser Ile His His Gly Val Val Glu Val Asp Ala Ala Val Thr 445 450 455 460	1575	
15 CCA GAG GAG CGC CAC CTG TCC AAG ATG CAG CAG AAC GGC TAC GAA AAT Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn 465 470 475	1623	
20 CCA ACC TAC AAG TTC TTT GAG CAG ATG CAG AAC TAT GGG GGC TTC ATG Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn Tyr Gly Gly Phe Met 480 485 490	1671	
TAGGATCCAT ATATAGGGCC CGGGTTATAA TTACCTCAGG TCGACCTAGA		1721

## (2) INFORMATION FOR SEQ ID NO:2:

25 (i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 492 amino acids	
(B) TYPE: amino acid	
(D) TOPOLOGY: linear	
30 (ii) MOLECULE TYPE: protein	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:	
Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ser Arg 1 5 10 15	
35 Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro 20 25 30	
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln 35 40 45	
40 Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp 50 55 60	
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu 65 70 75 80	

45

50

55

	Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn			
	85	90	95	
5	Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val			
	100	105	110	
	Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu			
	115	120	125	
10	Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys			
	130	135	140	
	Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu			
15	145.	150	155	160
	Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile			
	165	170	175	
20	Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu			
	180	185	190	
	Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val			
	195	200	205	
25	Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Glu Lys			
	210	215	220	
	Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu			
	225	230	235	240
30	Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu			
	245	250	255	
	Glu Ala Glu Glu Pro Tyr Glu Glu Ala Arg Glu Arg Thr Thr Ser Ile			
	260	265	270	
35	Ala Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Val Val Val Arg			
	275	280	285	
	Glu Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys Arg Ala Met Ile			
40	290	295	300	
	Ser Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe Phe			
	305	310	315	320
45	Tyr Gly Gly Cys Gly Gly Asn Arg Asn Asn Phe Asp Arg Glu Glu Tyr			
	325	330	335	
	Cys Met Ala Val Cys Gly Ser Ala Ile Pro Thr Thr Ala Ala Ser Thr			
	340	345	350	
50	Pro Asp Ala Val Asp Lys Tyr Leu Glu Arg Pro Lys Pro Gln Gln Phe			
	355	360	365	

Phe Gly Leu Met Gly Ser Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser  
 370 375 380

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val  
 5 385 390 395 400

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys  
 405 410 415

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val  
 10 420 425 430

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys His Tyr Thr Ser Ile  
 435 440 445

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg  
 15 450 455 460

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys  
 465 470 475 480

Phe Phe Glu Gln Met Gln Asn Tyr Gly Gly Phe Met  
 20 485 490

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3353 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGTTTCCTCG GCAGCGGTAG GCGAGAGCAC GCGGAGGAGC GTGCGCGGGG CCCCCGGGAGA	60
CGGCGGGCGGT GGCAGCGCGG GCAGAGCAAG GACCGCGCGG ATCCCACTCG CACAGCAGCG	120
CACTCGGTGC CCCGCGCAGG GTCGCGATGC TGCCCGGTTT GGCACTGCTC CTGCTGGCCG	180
CCTGGACGGC TCGGGCGCTG GAGGTACCCA CTGATGGTAA TGCTGGCCTG CTGGCTGAAC	240

45

50

55

	CCCAGATTGC CATGTTCTGT GGCAGACTGA ACATGCACAT GAATGTCCAG AATGGGAAGT	300
5	GGGATTCAAGA TCCATCAGGG ACCAAAACCT GCATTGATAAC CAAGGAAGGC ATCCTGCAGT	360
	ATTGCCAAGA AGTCTACCCCT GAACTGCAGA TCACCAATGT GGTAGAACGCC AACCAACCAG	420
	TGACCATCCA GAACTGGTGC AAGCGGGGCC GCAAGCAGTG CAAGACCCAT CCCCCACTTG	480
10	TGATTCCCTA CCGCTGCTTA GTTGGTGAGT TTGTAAGTGA TGCCCTTCTC GTTCCTGACA	540
	AGTGCAAATT CTTACACCAAG GAGAGGATGG ATGTTTGCAG AACTCATCTT CACTGGCACA	600
	CCGTCGCCAA AGAGACATGC AGTGAGAAGA GTACCAACTT GCATGACTAC GGCATGTTGC	660
	TGCCCTGCGG ATTGACAAG TTCCGAGGGG TAGAGTTGT GTGTTGCCA CTGGCTGAAG	720
15	AAAGTGACAA TGTGGATTCT GCTGATGCGG AGGAGGATGA CTCGGATGTC TGGTGGGGCG	780
	GAGCAGACAC AGACTATGCA GATGGGAGTG AAGACAAAGT AGTAGAAGTA GCAGAGGAGG	840
	AAGAAGTGGC TGAGGTGGAA GAAGAAGAAG CCGATGATGA CGAGGACGAT GAGGATGGTG	900
20	ATGAGGTAGA GGAAGAGGCT GAGGAACCCCT ACGAAGAACG CACAGAGAGA ACCACCAGCA	960
	TTGCCACAC CACCACCAAC ACCACAGAGT CTGTGGAAGA GGTGGTTCGA GTTCCTACAA	1020
	CAGCAGCCAG TACCCCTGAT GCCGTTGACA AGTATCTCGA GACACCTGGG GATGAGAATG	1080
25	AACATGCCA TTTCCAGAAA GCCAAAGAGA GGCTTGAGGC CAAGCACCGA GAGAGAATGT	1140
	CCCAGGTCA GAGAGAATGG GAAGAGGCAG AACGTCAAGC AAAGAACTTG CCTAAAGCTG	1200
	ATAAGAAGGC AGTTATCCAG CATTCCAGG AGAAAGTGG AATCTTGAA CAGGAAGCAG	1260
30	CCAACGAGAG ACAGCAGCTG GTGGAGACAC ACATGCCAG AGTGGAAAGCC ATGCTCAATG	1320
	ACCGCCGCCG CCTGGCCCTG GAGAACTACA TCACCGCTCT GCAGGCTGTT CCTCCTCGGC	1380
	CTCGTCACGT GTTCAATATG CTAAGAAAGT ATGTCCGCGC AGAACAGAAG GACAGACAGC	1440
35	ACACOCTAAA GCATTTCGAG CATGTGCGCA TGGTGGATCC CAAGAAAGCC GCTCAGATCC	1500
	GGTCCCAGGT TATGACACAC CTCCGTGTGA TTTATGAGCG CATGAATCAG TCTCTCTCCC	1560
	TGCTCTACAA CGTGCCTGCA GTGGCCGAGG AGATTCAAGGA TGAAGTTGAT GAGCTGCTTC	1620
	AGAAAGAGCA AAACTATTCA GATGACGTCT TGGCAACAT GATTAGTGAA CCAAGGATCA	1680
40	GTTACGGAAA CGATGCTCTC ATGCCATCTT TGACCGAAAC GAAAACCACC GTGGAGCTCC	1740
	TTCCCGTGAA TGGAGAGTTC AGCCTGGACG ATCTCCAGCC GTGGCATTCT TTTGGGGCTG	1800

45

50

55

	ACTCTGTGCC	AGCCAACACA	GAAAACGAAG	TTGAGCCTGT	TGATGCCCGC	CCTGCTGCCG	1860
	ACCGAGGACT	GACCACTCGA	CCAGGTTCTG	GGTTGACAAA	TATCAAGACG	GAGGAGATCT	1920
5	CTGAAGTGAA	GATGGATGCA	GAATTCCGAC	ATGACTCAGG	ATATGAAGTT	CATCATCAA	1980
	AATTGGTGT	CTTGCAGAA	GATGTGGT	CAAACAAAGG	TGCAATCATT	GGACTCATGG	2040
	TGGGCGGTGT	TGTCATAGCG	ACAGTGATCG	TCATCACCTT	GGTGATGCTG	AAGAAGAAC	2100
10	AGTACACATC	CATTCCATCAT	GGTGTGGTGG	AGGTTGACGC	CGCTGTCACC	CCAGAGGAGC	2160
	GCCACCTGTC	CAAGATGCAG	CAGAACGGCT	ACGAAAATCC	AACCTAGAAG	TTCTTGAGC	2220
	AGATGCAGAA	CTAGACCCCC	GCCACAGCAG	CCTCTGAAGT	TGGACAGCAA	AACCATTGCT	2280
15	TCACTACCCA	TCGGTGTCCA	TTTATAGAAT	AATGTGGAA	GAAACAAACC	CGTTTATGA	2340
	TTTACTCATT	ATCGCCTTTT	GACAGCTGTG	CTGTAACACA	AGTAGATGCC	TGAACTTGAA	2400
	TTAATCCACA	CATCAGTAAT	GTATTCTATC	TCTCTTACA	TTTGGTCTC	TATACTACAT	2460
	TATTAATGGG	TTTGTGTAC	TGTAAAGAAT	TTAGCTGTAT	CAAACACTGT	CATGAATAGA	2520
20	TTCTCTCCTG	ATTATTTATC	ACATAGCCCC	TTAGCCAGTT	GTATATTATT	CTTGTGGTTT	2580
	GTGACCCAAT	TAAGTCCTAC	TTTACATATG	CTTTAAGAAT	CGATGGGGGA	TGCTTCATGT	2640
	GAACGTGGGA	GTTCAAGCTGC	TTCTCTTGCC	TAAGTATTCC	TTTCTGATC	ACTATGCATT	2700
25	TTAAAGTTAA	ACATTTTAA	GTATTCAGA	TGCTTAGAG	AGATTTTTT	TCCATGACTG	2760
	CATTTACTG	TACAGATTGC	TGCTTCTGCT	ATATTTGTGA	TATAGGAATT	AAGAGGATAC	2820
	ACACGTTGT	TTCTTCGTGC	CTGTTTATG	TGCACACATT	AGGCATTGAG	ACTTCAAGCT	2880
30	TTTCTTTTTT	TGTCCACGTA	TCTTGGTGT	TTTGATAAAG	AAAAGAATCC	CTGTTCATG	2940
	TAAGCACTT	TACGGGGCGG	GTGGGGAGGG	GTGCTCTGCT	GGTCTTCAAT	TACCAAGAAT	3000
	TCTCCAAAAC	AATTTCTGC	AGGATGATTG	TACAGAATCA	TTGCTTATGA	CATGATCGCT	3060
35	TTCTACACTG	TATTAGATAA	ATAAATTTAA	AAAAATAACC	CCGGGCAAGA	CTTTCTTTG	3120
	AAGGATGACT	ACAGACATTA	AATAATCGAA	GTAATTTGG	GTGGGGAGAA	GAGGCAGATT	3180
	CAATTTCTT	TAACCAGTCT	GAAGTTCAT	TTATGATACA	AAAGAAGATG	AAAATGGAAG	3240
40	TGGCAATATA	AGGGGATGAG	GAAGGCATGC	CTGGACAAAC	CCTTCTTTA	AGATGTGTCT	3300
	TCAATTGTA	AAAATGGTG	TTTCATGTA	AATAAATACA	TTCTTGAGG	AGC	3353

45

50

55

## (2) INFORMATION FOR SEQ ID NO:4:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: peptide

10 (iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

15 (vi) ORIGINAL SOURCE:

15 (A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

20 Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys  
1 5 10 15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile  
20 25 30

25 Gly Leu Met Val Gly Val Val Ile Ala  
35 40

## (2) INFORMATION FOR SEQ ID NO:5:

- 30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: peptide

35 (iii) HYPOTHETICAL: NO

40 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

45 Arg Pro Lys Pro Gln Gln Phe Phe Gly Leu Met  
1 5 10

## (2) INFORMATION FOR SEQ ID NO:6:

- 5           (i) SEQUENCE CHARACTERISTICS:  
               (A) LENGTH: 21 base pairs  
               (B) TYPE: nucleic acid  
               (C) STRANDEDNESS: single  
               (D) TOPOLOGY: linear

10           (ii) MOLECULE TYPE: DNA (genomic)

15           (iii) HYPOTHETICAL: NO

20           (iv) ANTI-SENSE: NO

25           (ix) FEATURE:

- 25            (A) NAME/KEY: CDS  
               (B) LOCATION: 1..21

30           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

20           CAT CAA AAA TTG GTG TTC TTT  
          His Gln Lys Leu Val Phe Phe  
           1               5

21

## 25           (2) INFORMATION FOR SEQ ID NO:7:

- 30           (i) SEQUENCE CHARACTERISTICS:  
               (A) LENGTH: 7 amino acids  
               (B) TYPE: amino acid  
               (D) TOPOLOGY: linear

35           (ii) MOLECULE TYPE: protein

35           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

35           His Gln Lys Leu Val Phe Phe  
           1               5

## 35           (2) INFORMATION FOR SEQ ID NO:8:

- 40           (i) SEQUENCE CHARACTERISTICS:  
               (A) LENGTH: 21 base pairs  
               (B) TYPE: nucleic acid  
               (C) STRANDEDNESS: single  
               (D) TOPOLOGY: linear

45           (ii) MOLECULE TYPE: DNA (genomic)

45           (iii) HYPOTHETICAL: NO

50

55

(iv) ANTI-SENSE: NO

5 (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

10 CAT CAA GAG TTG GTG TTC TTT  
His Gln Glu Leu Val Phe Phe  
1 5

21

15 (2) INFORMATION FOR SEQ ID NO:9:

- 15 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

His Gln Glu Leu Val Phe Phe  
1 5

25 (2) INFORMATION FOR SEQ ID NO:10:

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

35 (iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

45

50

55

CAT CAA GTG TTG GTG TTC TTT  
 His Gln Val Leu Val Phe Phe  
 1 5

21

5

## (2) INFORMATION FOR SEQ ID NO:11:

10

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

15

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

20

His Gln Val Leu Val Phe Phe  
 1 5

## (2) INFORMATION FOR SEQ ID NO:12:

25

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

30

- (ii) MOLECULE TYPE: DNA (genomic)

35

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

40

- (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..21

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

45

CAT CAA GTG TTG GTG TTC TTT  
 His Gln Val Leu Val Phe Phe  
 1 5

21

## (2) INFORMATION FOR SEQ ID NO:13:

50

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

55

50

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

5 His Gln Val Leu Val Phe Phe  
1 5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20 (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

25 CAT CAA AAA TTG GTG CCG TTT  
His Gln Lys Leu Val Pro Phe  
1 5

21

30 (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

His Gln Lys Leu Val Pro Phe  
1 5

40 (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid

45

50

55

(C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 1..27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

15 GTC ATA GCG ACA GTG ATC GTC ATC ACC  
 Val Ile Ala Thr Val Ile Val Ile Thr  
     1                     5

27

20 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Val Ile Ala Thr Val Ile Val Ile Thr  
     1                     5

30 (2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

35 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..27

45

50

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GTC ATA GCG ACA GTG ATC ATC ATC ACC	27
Val Ile Ala Thr Val Ile Ile Ile Thr	
5 1 5	

(2) INFORMATION FOR SEQ ID NO:19:

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Val Ile Ala Thr Val Ile Ile Ile Thr	
1 5	

(2) INFORMATION FOR SEQ ID NO:20:

- 20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 30 (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..27

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GTC ATA GCG ACA GTG ATC GGC ATC ACC	27
Val Ile Ala Thr Val Ile Gly Ile Thr	
1 5	

40 (2) INFORMATION FOR SEQ ID NO:21:

45

50

55

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Val Ile Ala Thr Val Ile Gly Ile Thr  
 10 1 5

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- 25 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

30 GTC ATA GCG ACA GTG ATC TTC ATC ACC  
 Val Ile Ala Thr Val Ile Phe Ile Thr  
 1 5

27

(2) INFORMATION FOR SEQ ID NO:23:

- 35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Val Ile Ala Thr Val Ile Phe Ile Thr  
 1 5

45

50

55

## (2) INFORMATION FOR SEQ ID NO:24:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

15 (ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 1..27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

20 CAA AAA TTG GTG TTC TTT GCG GAA GAT  
 Gln Lys Leu Val Phe Phe Ala Glu Asp  
 1 5

27

## (2) INFORMATION FOR SEQ ID NO:25:

- 25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Gln Lys Leu Val Phe Phe Ala Glu Asp  
 1 5

## 35 (2) INFORMATION FOR SEQ ID NO:26:

- 40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

45

50

55

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- 5 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

10 CAA AAA TTG GTG TTC TTT GCA CAA GAT  
 Gln Lys Leu Val Phe Phe Ala Gln Asp  
 1 5

27

(2) INFORMATION FOR SEQ ID NO:27:

- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Gln Lys Leu Val Phe Phe Ala Gln Asp  
 1 5

25

### Claims

1. A nucleic acid molecule encoding an amyloid precursor mutein, wherein the nucleic acid molecule comprises, from the 5' end to the 3' end a nucleic acid sequence encoding a marker and a nucleic acid sequence encoding amino acid up to but excluding nucleic acids encoding BAP domain.
2. A nucleic acid molecule encoding an amyloid precursor mutein, wherein the nucleic acid molecule comprises, from the 5' end to the 3' end a nucleic acid sequence encoding BAP domain and a nucleic acid sequence encoding a marker.
3. A nucleic acid molecule which comprises the nucleic acid molecule of claim 1 ligated to the nucleic acid molecule of claim 2.
4. The nucleic acid molecule of claim 3, wherein the nucleic acid molecule is a nucleic acid molecule selected from the group consisting of DNA, cDNA or RNA.
5. The nucleic acid molecule of claim 1, wherein the nucleic acid molecule is selected from the group consisting of pCLL983, pCLL935, pCLL934 and pCLL913.
6. The nucleic acid molecule of claim 2, wherein the nucleic acid molecule is selected from the group consisting of pCLL947, pCLL914, pCLL937, pCLL949 and pCLL957.
7. The nucleic acid molecule of claim 3, wherein the nucleic acid molecule is selected from the group consisting of pCLL619, pCLL620, pCLL618, pCLL964, pCLL962, pCLL989, pCLL987, pCLL990, pCLL988, pCLL600, pCLL601, pCLL602, pCLL603, pCLL604, pCLL605, pCLL606 and pCLL607.
8. A vector comprising the nucleic acid molecule of claim 1, claim 2 or claim 3.
9. A cell comprising the nucleic acid molecule of claim 1, claim 2 or claim 3.
10. A recombinant polypeptide produced by the nucleic acid molecule of claim 1, produced by the nucleic acid molecule of claim 2 or produced by the nucleic acid molecule of claim 3.

11. A method of detecting the presence of the recombinant polypeptides of claim 10 in a sample, which comprises contacting an antibody directed to the marker and the sample under suitable conditions to favor the formation of an antibody-antigen complex, and detecting the presence of any complex so formed.

5

10

15

20

25

30

35

40

45

50

55

Figure 1.

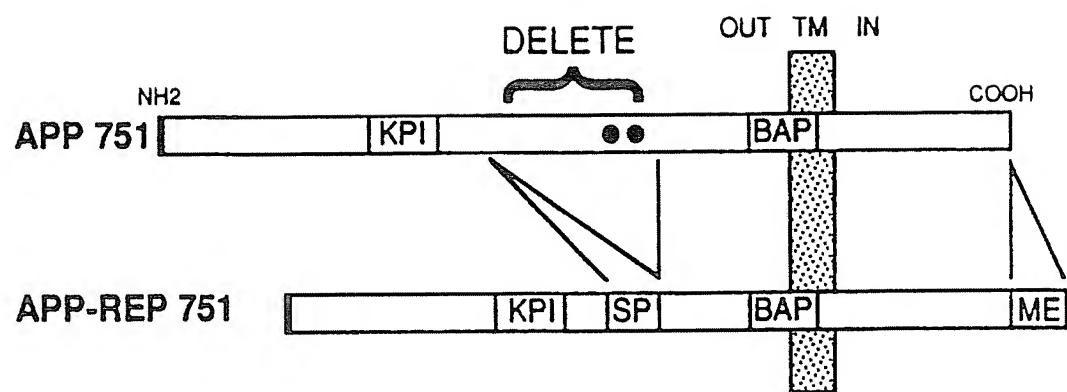


Figure 2.

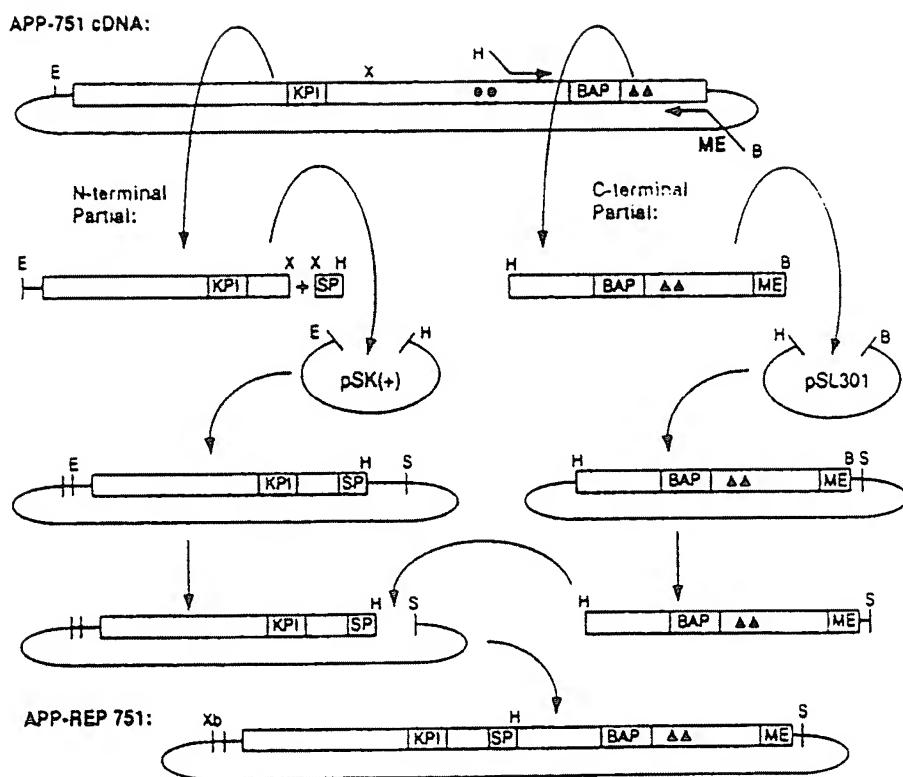


Figure 3.

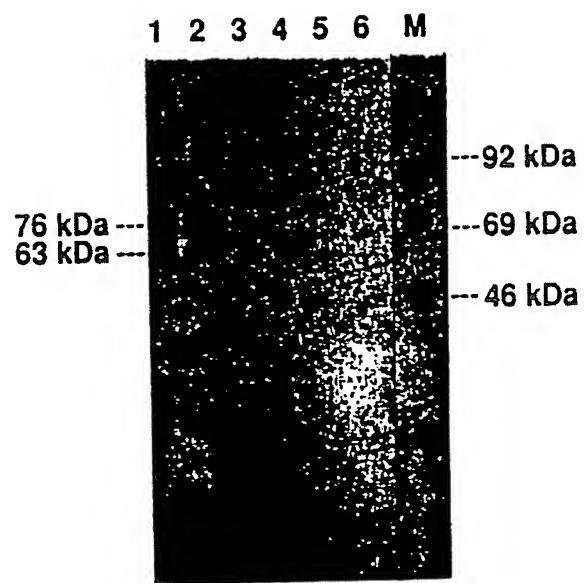


Figure 4.

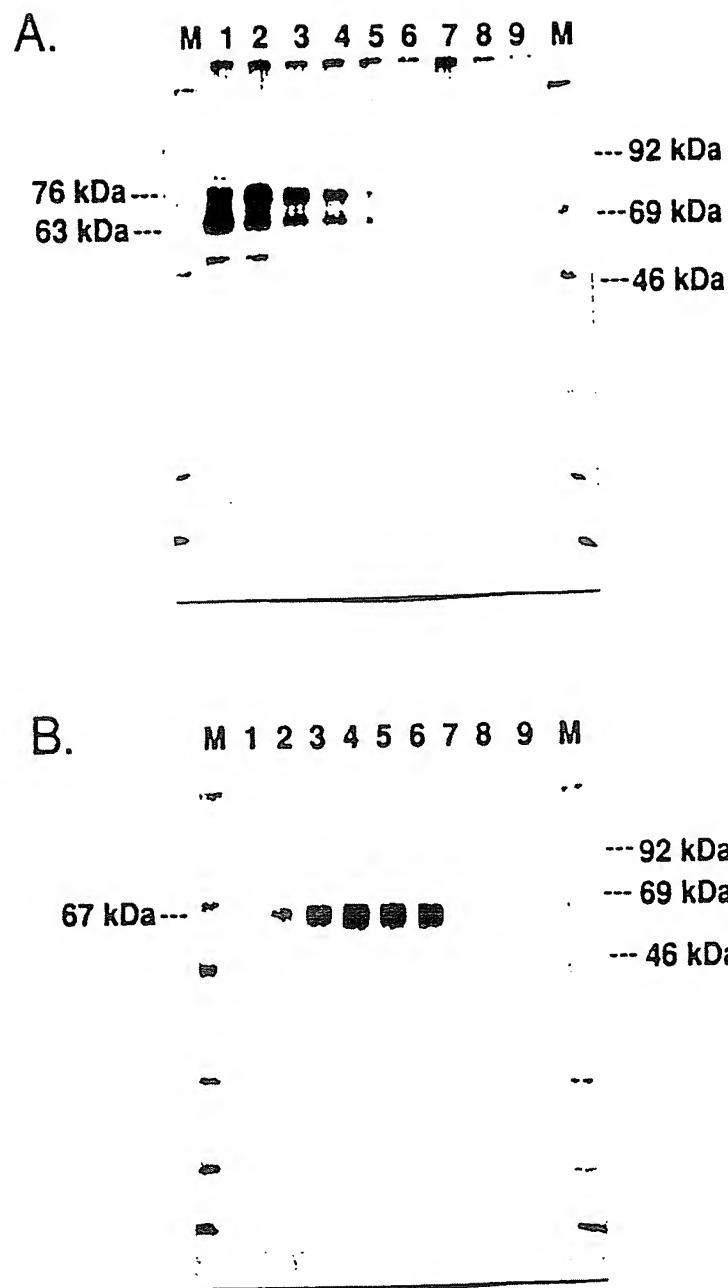
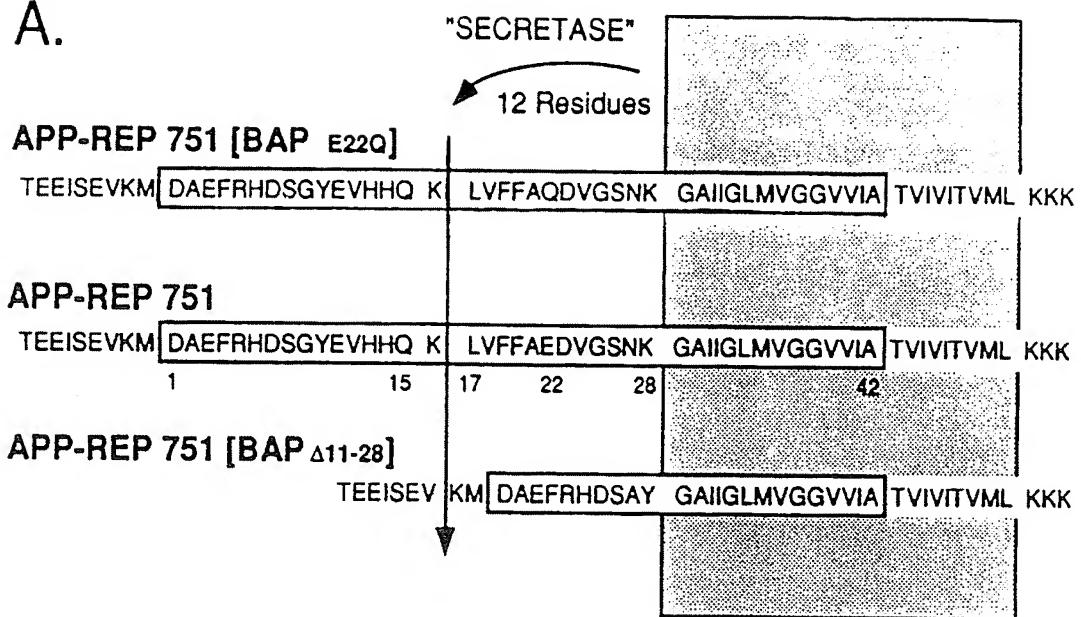


Figure 5.

A.



B.

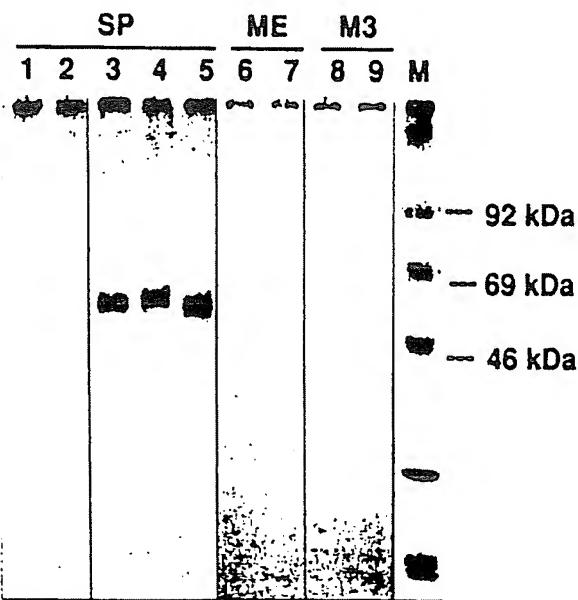
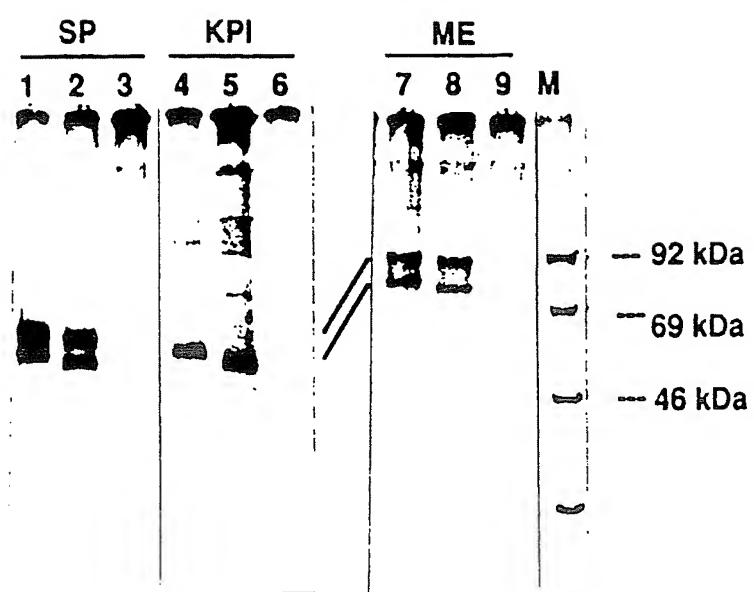


Figure 5.

C.



D.

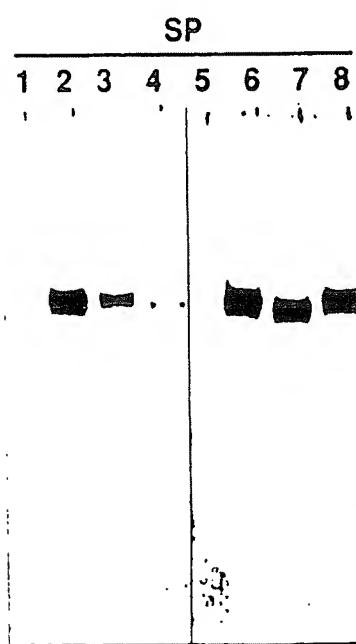


Figure 5.

E.

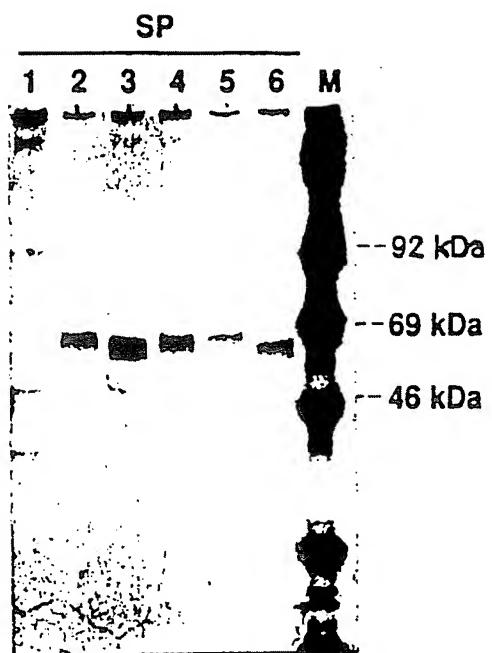
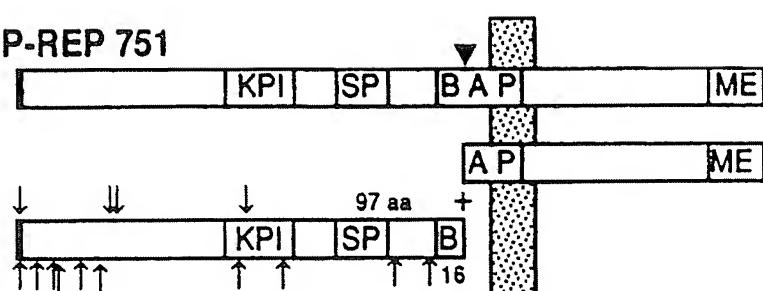


Figure 6.

A.

APP-REP 751



APP-REP BAP  $\Delta$ 11-28

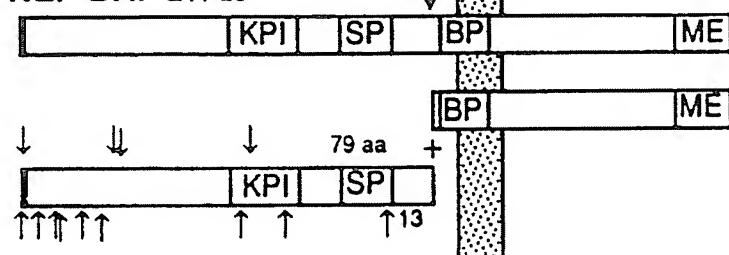


FIGURE 6.

B.

1 2 3

21.5 kDa



SEQUENCE: pCLL602 (APP-REP 751 protein)  
 VECTOR: pcDNA-I-neo (Invitrogen)  
 pcDNA-I-neo-XS (JSJ modified polylinker to permit directional  
 subcloning into XbaI-SalI sites)  
 INSERT: XbaI-SalI fragment encoding APP-REP from pCLL964 16-1711  
 SEQUENCE: 5' polylinker:  
     HindIII-XbaI from pcDNA-I-neo-XS 2-15  
     XbaI-EcoRI from pBluescript SK+ 16-47  
 APP-REP 751:  
     Amino-terminal partial from pCLL935): 48-1314  
         5' untranslated APP cDNA (from EcoRI) 48-195  
         N-terminal APP (to XhoI) 196-1273  
         Substance P marker (XhoI to HindIII) 1274-1314  
     Carboxy-terminal partial from pCLL947): 1314-1671  
         C-terminal APP and BAP (from novel HindIII) 1314-1696  
         Met-enkephalin marker (plus stop codon) 1657-1674  
 3' polylinker:  
     BamHI-SalI from pSL301 1674-1711  
     SalI-end of sequence from pcDNA-I-neo-XS 1712-1721

10	20	30	40	50
*	*	*	*	*
AAGCTTGGGG ATCCGCTCTA GAACTAGTGG ATCCCCCGGG CTGCAGGAAT				
TTCGAACCCC TAGGCGAGAT CTTGATCACC TAGGGGGCCC GACGTCTTA				
60	70	80	90	100
*	*	*	*	*
TCGGGGGGGG CAGCGGTAGG CGAGAGCACCG CGGAGGAGCG TGCGCGGGGC				
AGCCCCCCCTT GTGCCCATCC GCTCTCGTGC GCCTCCTCGC ACGCGCCCCG				
110	120	130	140	150
*	*	*	*	*
CCCAGGAGAC GGCGCGGGTG GCGGCGCGGG CAGAGCAAGG ACGCGGCGGA				
GGGCCCTCTG CGCCCGGCCAC CGCCGCGCCC GTCTCGTTCC TGCAGCGCCT				
160	170	180	190	
*	*	*	*	
TCCCACTCGC ACAGCAGCGC ACTCGGTGCC CGCGCGAGGG TCGCG				
AGGGTGAGCG TGTCGTCGCG TGAGCCACGG GGCGCGTCCC AGCGC				
200	210	220	230	240
*	*	*	*	*
ATG CTG CCC GGT TTG GCA CTG CTC CTG CTG GCC GCC TGG ACG GCT				
TAC GAC GGG CCA AAC CGT GAC GAG GAC GAC CGG CGG ACC TGC CGA				
Met Leu Pro Gly Leu Ala Leu Leu Ala Ala Ala Trp Thr Ala>				
250	260	270	280	
*	*	*	*	
CGG GCG CTG GAG GTA CCC ACT GAT GGT AAT GCT GGC CTG CTG GCT GAA				
GCC CGC GAC CTC CAT GGG TGA CTA CCA TTA CGA CCG GAC CGA CTT				
Arg Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu>				

290            300            310            320            330  
 \*            \*            \*            \*            \*  
 CCC CAG ATT GCC ATG TTC TGT GGC AGA CTG AAC ATG CAC ATG AAT GTC  
 GGG GTC TAA CGG TAC AAG ACA CCG TCT GAC TTG TAC GTG TAC TTA CAG  
 Pro Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val>  
  
 340            350            360            370            380  
 \*            \*            \*            \*            \*  
 CAG AAT GGG AAG TGG GAT TCA GAT CCA TCA GGG ACC AAA ACC TGC ATT  
 GTC TTA CCC TTC ACC CTA AGT CTA GGT AGT CCC TGG TTT TGG ACC TAA  
 Gln Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile>  
  
 390            400            410            420            430  
 \*            \*            \*            \*            \*  
 GAT ACC AAG GAA GGC ATC CTG CAG TAT TGC CAA GAA GTC TAC CCT GAA  
 CTA TGG TTC CTT CCG TAG GAC GTC ATA ACG GTT CTT CAG ATG GGA CTT  
 Asp Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu>  
  
 440            450            460            470            480  
 \*            \*            \*            \*            \*  
 CTG CAG ATC ACC AAT GTG GTA GAA GCC AAC CAA CCA GTG ACC ATC CAG  
 GAC GTC TAG TGG TTA CAC CAT CTT CCG TTG GTT GGT CAC TGG TAG GTC  
 Leu Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln>  
  
 490            500            510            520  
 \*            \*            \*            \*  
 AAC TGG TGC AAG CGG GGC CGC AAG CAG TGC AAG ACC CAT CCC CAC TTT  
 TTG ACC ACG TTC GCC CCG GCG TTC GTC ACG TTC TGG GTA GGG GTG AAA  
 Asn Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe>  
  
 530            540            550            560            570  
 \*            \*            \*            \*            \*  
 GTG ATT CCC TAC CGC TGC TTA GTT GGT GAG TTT GTA AGT GAT GGC CTT  
 CAC TAA GGG ATG GCG ACG AAT CAA CCA CTC AAA CAT TCA CTA CGG GAA  
 Val Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu>  
  
 580            590            600            610            620  
 \*            \*            \*            \*            \*  
 CTC GTT CCT GAC AAG TGC AAA TTC TTA CAC CAG GAG AGG ATG GAT GTT  
 GAG CAA GGA CTG TTC ACG TTT AAG AAT GTG GTC CTC TCC TAC CTA CAA  
 Leu Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val>  
  
 630            640            650            660            670  
 \*            \*            \*            \*            \*  
 TGC GAA ACT CAT CTT CAC TGG CAC ACC GTC GCC AAA GAG ACA TGC AGT  
 ACG CTT TGA GTA GAA GTG ACC GTG TGG CAG CGG TTT CTC TGT ACG TCA  
 Cys Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser>  
  
 680            690            700            710            720  
 \*            \*            \*            \*            \*  
 GAG AAG AGT ACC AAC TTG CAT GAC TAC GGC ATG TTG CTG CCC TGC GGA  
 CTC TTC TCA TGG TTG AAC GTA CTG ATG CCG TAC AAC GAC GGG ACG CCT  
 Glu Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly>

FIGURE 7  
(continued)

730            740            750            760

ATT GAC AAG TTC CGA GGG GTA GAG TTT GTG TGT TGC CCA CTG GCT GAA  
 TAA CTG TTC AAG GCT CCC CAT CTC AAA CAC ACA ACG GGT GAC CGA CTT  
 Ile Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu>

770            780            790            800            810

GAA AGT GAC AAT GTG GAT TCT GCT GAT GCG GAG GAG GAT GAC TCG GAT  
 CTT TCA CTG TTA CAC CTA AGA CGA CTA CGC CTC CTC CTA CTG AGC CTA  
 Glu Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp>

820            830            840            850            860

GTC TGG TGG GGC GGA GCA GAC ACA GAC TAT GCA GAT GGG AGT GAA GAC  
 CAG ACC ACC CCG CCT CGT CTG TGT CTG ATA CGT CTA CCC TCA CTT CTG  
 Val Trp Trp Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp>

870            880            890            900            910

AAA GTA GTA GAA GTA GCA GAG GAG GAA GAA GTG GCT GAG GTG GAA GAA  
 TTT CAT CAT CTT CAT CGT CTC CTC CTT CTT CAC CGA CTC CAC CTT CTT  
 Lys Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu>

920            930            940            950            960

GAA GAA GCC GAT GAT GAC GAG GAC GAT GAG GAT GGT GAT GAG GAA GAG  
 CTT CTT CGG CTA CTA CTG CTC CTG CTA CTC CTA CCA CTA CTC CAT CTC  
 Glu Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu>

970            980            990            1000

GAA GAG GCT GAG GAA CCC TAC GAA GAA GCC ACA GAG AGA ACC ACC AGC  
 CTT CTC CGA CTC CTT GGG ATG CTT CTT CGG TGT CTC TCT TGG TGG TCG  
 Glu Glu Ala Glu Glu Pro Tyr Glu Ala Thr Glu Arg Thr Thr Ser>

1010           1020           1030           1040           1050

ATT GCC ACC ACC ACC ACC ACC ACA GAG TCT GTG GAA GAG GTG GTT  
 TAA CGG TGG TGG TGG TGG TGG TGT CTC AGA CAC CTT CTC CAC CAA  
 Ile Ala Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Val Val>

1060           1070           1080           1090           1100

CGA GAG GTG TGC TCT GAA CAA GCC GAG ACG GGG CCG TGC CGA GCA ATG  
 GCT CTC CAC ACG AGA CTT GTT CGG CTC TGC CCC GGC ACG GCT CGT TAC  
 Arg Glu Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys Arg Ala Met>

1110           1120           1130           1140           1150

ATC TCC CGC TGG TAC TTT GAT GTG ACT GAA GGG AAG TGT GCC CCA TTC  
 TAG AGG GCG ACC ATG AAA CTA CAC TGA CTT CCC TTC ACA CGG GGT AAG  
 Ile Ser Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe>

FIGURE 7  
(continued)

1160	1170	1180	1190	1200
*	*	*	*	*
TTT TAC GGC GGA TGT GGC GGC AAC CGG AAC AAC TTT GAC ACA GAA GAG AAA ATG CCG CCT ACA CCG CCG TTG GCC TTG AAA CTG TGT CTT CTC Phe Tyr Gly Cys Gly Gly Asn Arg Asn Asn Phe Asp Thr Glu Glu>				
1210	1220	1230	1240	
*	*	*	*	
TAC TGC ATG GCC GTG TGT GGC AGC GCC ATT CCT ACA ACA GCA GCC AGT ATG ACG TAC CGG CAC ACA CCG TCG CGG TAA GGA TGT TGT CGT CGG TCA Tyr Cys Met Ala Val Cys Gly Ser Ala Ile Pro Thr Thr Ala Ala Ser>				
1250	1260	1270	1280	1290
*	*	*	*	*
ACC CCT GAT GCC GTT GAC AAG TAT CTC GAG CGG CCC AAG CCC CAG CAG TGG GGA CTA CGG CAA CTG TTC ATA GAG CTC GCC GGG TTC GGG GTC GTC Thr Pro Asp Ala Val Asp Lys Tyr Leu Glu Arg Pro Lys Pro Gln Gln>				
1300	1310	1320	1330	1340
*	*	*	*	*
TTC TTT GGC CTG ATG GGA AGC TTG ACA AAT ATC AAG ACG GAG GAG ATC AAG AAA CCG GAC TAC CCT TCG AAC TGT TTA TAG TTC TGC CTC CTC TAG Phe Phe Gly Leu Met Gly Ser Leu Thr Asn Ile Lys Thr Glu Glu Ile>				
1350	1360	1370	1380	1390
*	*	*	*	*
TCT GAA GTG AAG ATG GAT GCA GAA TTC CGA CAT GAC TCA GGA TAT GAA AGA CTT CAC TTC TAC CTA CGT CTT AAG GCT GTC CTG AGT CCT ATA CTT Ser Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu>				
1400	1410	1420	1430	1440
*	*	*	*	*
GTT CAT CAT CAA AAA TTG GTG TTC TTT GCA GAA GAT GTG GGT TCA AAC CAA GTA GTA GTT TTT AAC CAC AAG AAA CGT CTT CTA CAC CCA AGT TTG Val His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn>				
1450	1460	1470	1480	
*	*	*	*	
AAA GGT GCA ATC ATT GGA CTC ATG GTG GGC GGT GTT GTC ATA GCG ACA TTT CCA CGT TAG TAA CCT GAG TAC CAC CCG CCA CAA CAG TAT CGC TGT Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr>				
1490	1500	1510	1520	1530
*	*	*	*	*
GTG ATC GTC ATC ACC TTG GTG ATG CTG AAG AAG AAA CAG TAC ACA TCC CAC TAG CAG TAG TGG AAC CAC TAC GAC TTC TTC TTT GTC ATG TGT AGG Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser>				
1540	1550	1560	1570	1580
*	*	*	*	*
ATT CAT CAT GGT GTG GTG GAG GTT GAC GCC GCT GTC ACC CCA GAG GAG TAA GTA GTA CCA CAC CAC CTC CAA CTG CGG CGA CAG TGG GGT CTC CTC Ile His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu>				

FIGURE 7  
(continued)

30            1600            1610            1620            1630  
\*            \*            \*            \*            \*  
CGC CAC CTG TCC AAG ATG CAG CAG AAC GGC TAC GAA AAT CCA ACC TAC  
GCG GTG GAC AGG TTC TAC GTC GTC TTG CCG ATG CTT TTA GGT TGG ATG  
Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr>  
  
1640            1650            1660            1670            1680  
\*            \*            \*            \*            \*  
AAG TTC TTT GAG CAG ATG CAG AAC TAT GGG GGC TTC ATG TAG GATCCA  
TTC AAG AAA CTC GTC TAC GTC TTG ATA CCC CCG AAG TAC ATC CTAGGT  
Lys Phe Phe Glu Gln Met Gln Asn Tyr Gly Gly Phe Met \*\*\*  
  
1690            1700            1710            1720  
\*            \*            \*            \*  
TATATAGGGC CCGGGTTAT ATTACCTCA GGTGACCTA GA  
ATATATCCCG GGCCCAATA TTAATGGAGT CCAGCTGGAT CT

FIGURE 7  
(continued)

Total of 12 files.  
 PRMRD3>TYPE APP770.  
 LCCUS HUMAFPA4 3353 bp ss-mRNA PRI 15-JUN-1989  
 Human amyloid A4 mRNA, complete cds.  
 ACCESSION Y00264  
 KEYWORDS amyloid fibril protein; cell surface glycoprotein.  
 SOURCE human (Homo sapiens).  
 ORGANISM Homo sapiens  
 Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
 Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.  
 REFERENCE 1 (bases 1 to 3353; enum. -146 to 3207, no zero)  
 AUTHORS Kang,J., Lemaire,H.G., Unterbeck,A., Salbaum,J.M., Masters,C.L.,  
 Grzeschik,K.-H., Multhaup,G., Beyreuther,K. and Mueller-Hill,B.  
 TITLE The precursor of Alzheimer's disease amyloid A4 protein resembles a  
 cell-surface receptor  
 JOURNAL Nature 325, 733-736 (1987)  
 STANDARD simple automatic  
 REFERENCE 2 (bases 1 to 3353; enum. 1 to 3353)  
 AUTHORS Mueller Hill,B.  
 JOURNAL Unpublished (1987) Submitted to the EMBL data library.  
 STANDARD simple automatic  
 COMMENT \*source: tissue=cortex of brain; \*source: developmental  
 stage=5-month-old aborted fetus;  
 EMBL features not translated to GenBank features:  
 key from to description  
 SITE 3080 3085 polyA signal  
 SITE 3089 3094 polyA signal  
 SITE 3331 3336 polyA signal  
 POLYA 3353 3353 polyA site  
 FEATURES from to/span description  
 pept 147 2234 amyloid A4 /nomgen="APP" /map="21q21.2"  
 /hgml locus uid="LG0136J"  
 BASE COUNT 922 a 745 c 867 g 819 t  
 ORIGIN

App770. Length: 3353 May 29, 1990 11:30 Check: 6510 ..

```

1 AGTTTCCTCG GCAGCGGTAG GCGAGAGCAC GCGGAGGAGC GTGCCGGGG
51 CCCCCGGGAGA CGGCGGCGGT GGCGGCGCGG GCAGAGCAAG GACGCCGG
101 ATCCCACCTCG CACAGCAGCG CACTCGGTGC CCCGCGCAGG GTCGGATGC
151 TGCCCCGGTTT GGCACTGCTC CTGCTGGCCCG CCTGGACGGC TCGGGCGCTG
201 GAGGTACCCA CTGATGGTAA TGCTGGCCTG CTGGCTGAAC CCCAGATTGC
251 CATGTTCTGT GGCAGACTGA ACATGCACAT GAATGTCCAG AATGGGAAGT
301 GGGATTCAAGA TCCATCAGGG ACCAAAACCT GCATTGATAC CAAGGAAGGC
351 ATCCTGCAGT ATTGCCAAGA AGTCTACCCT GAACTGCAGA TCACCAATGT
401 GGTAGAACGCC AACCAACCAG TGACCATCCA GAACTGGTGC AAGCGGGGCC
451 GCAAGCAGTG CAAGACCCAT CCCCCACTTTC TGATTCCCTA CCGCTGCTTA
501 GTTGGTGAGT TTGTAAGTGA TGCCCTTCTC GTTCCTGACA AGTGCAAATT
551 CTTACACCAAG GAGAGGATGG ATGTTGCGA AACTCATCTT CACTGGCACA
601 CCGTCGCCAA AGAGACATGC AGTGAGAAGA GTACCAAACCT GCATGACTAC
651 GGCATGTTGC TGCCCTGCAG AATTGACAAG TTCCGAGGGG TAGAGTTGT
701 GTGTTGCCCA CTGGCTGAAG AAAGTGACAA TGTGGATTCT GCTGATGCCG

```

FIGURE 8

751 AGGAGGGATGA CTCGGA<sub>13</sub>TC TGGTGGGGCG GAGCAGACAC AGA<sub>1</sub> TATGCA  
 801 GATGGGAGTG AAGACAAAGT AGTAGAAGTA GCAGAGGAGG AAGAAGTGGC  
 851 TGAGGTGGAA GAAGAAGAAG CCGATGATGA CGAGGACGAT GAGGATGGTG  
 901 ATGAGGTAGA GGAAGAGGCT GAGGAACCCCT ACGAAGAACGC CACAGAGAGA  
 951 ACCACCAGCA TTGCCACAC CACCACCAAC ACCACAGAGT CTGTGGAAGA  
 1001 GGTGGTTCGA GTTCTACAA CAGCAGCCAG TACCCCTGAT GCCGTTGACA  
 1051 AGTATCTCGA GACACCTGGG GATGAGAATG AACATGCCA TTTCCAGAAA  
 1101 GCCAAAGAGA GGCTTGAGGC CAAGCACCGA GAGAGAATGT CCCAGGTCT  
 1151 GAGAGAATGG GAAGAGGCAG AACGTCAAGC AAAGAACTTG CCTAAAGCTG  
 1201 ATAAGAAGGC AGTTATCCAG CATTTCCAGG AGAAAGTGGA ATCTTTGGAA  
 1251 CAGGAAGCAG CCAACGAGAG ACAGCAGCTG GTGGAGACAC ACATGGCCAG  
 1301 AGTGGAAAGCC ATGCTCAATG ACCGCCGCCG CCTGGCCCTG GAGAACTACA  
 1351 TCACCGCTCT GCAGGCTGTT CCTCCTCGGC CTCGTCACGT GTTCAATATG  
 1401 CTAAAGAAGT ATGTCCGCGC AGAACAGAAG GACAGACAGC ACACCCCTAAA  
 1451 GCATTTCGAG CATGTGCGCA TGGTGGATCC CAAGAAAGCC GCTCAGATCC  
 1501 GGTCCCAGGT TATGACACAC CTCCGTGTGA TTTATGAGCG CATGAATCAG  
 1551 TCTCTCTCCC TGCTCTACAA CGTGCCTGCA GTGGCCGAGG AGATTCAAGGA  
 1601 TGAAAGTTGAT GAGCTGCTTC AGAAAGAGCA AAACATATTCA GATGACGTCT  
 1651 TGGCCAACAT GATTAGTGAA CCAAGGATCA GTTACGGAAA CGATGCTCTC  
 1701 ATGCCATCTT TGACCGAAAC GAAAACCACC GTGGAGCTCC TTCCCCTGAA  
 1751 TGGAGAGTTC AGCCTGGACG ATCTCCAGCC GTGGCATTCT TTTGGGGCTG  
 1801 ACTCTGTGCC AGCCAACACA GAAAACGAAG TTGAGCCTGT TGATGCCCGC  
 1851 CCTGCTGCCG ACCGAGGACT GACCACTCGA CCAGGTTCTG GGTGACAAA  
 1901 TATCAAGACG GAGGAGATCT CTGAAGTGAA GATGGATGCA GAATTCCGAC  
 1951 ATGACTCAGG ATATGAAGTT CATCATCAA AATTGGTGTT CTTTGAGAA  
 2001 GATGTGGGTT CAAACAAAGG TGCAATCATT GGACTCATGG TGGCGGTGT  
 2051 TGTCAAGACG AGAGTGTGTC TCATCACCTT GGTGATGCTG AAGAAGAAC  
 2101 AGTACACATC CATTGATCAT GGTGTGGTGG AGGTTGACGC CGCTGTCACC  
 2151 CCAGAGGAGC GCCACCTGTC CAAGATGCAG CAGAACGGCT ACGAAAATCC  
 2201 AACCTACAAG TTCTTTGAGC AGATGCAGAA CT

AGACCCCC GCCACAGCAG

2251 CCTCTGAAGT TGGACAGCAA AACCAATTGCT TCACTACCCA TCGGTGTCCA

2301 TTTATAGAAT AATGTG<sup>CC</sup>AA GAAACAAACC CGTTTATGA TTT<sup>A</sup>CTCATT  
 2351 ATCGCCTTT GACAGCTGTG CTGTAACACA AGTAGATGCC TGA<sup>A</sup>CTTGAA  
 2401 TTAATCCACA CATCAGTAAT GTATTCTATC TCTCTTACA TTTGGTCTC  
 2451 TATACTACAT TATTAATGGG TTTTGTGTAC TGTAAAGAAT TTAGCTGTAT  
 2501 CAAACTAGTG CATGAATAGA TTCTCTCCTG ATTATTATC ACATAGCCCC  
 2551 TTAGCCAGTT GTATATTATT CTTGTGGTT GTGACCCAAT TAAGTCCTAC  
 2601 TTTACATATG CTTAAAGAAT CGATGGGGGA TGCTTCATGT GAACGTGGGA  
 2651 GTTCAGCTGC TTCTCTGCC TAAGTATTCC TTTCCTGATC ACTATGCATT  
 2701 TTAAAGTTAA ACATTTTAA GTATTCAGA TGCTTAGAG AGATTTTTT  
 2751 TCCATGACTG CATTACTG TACAGATTGC TGCTCTGCT ATATTTGTGA  
 2801 TATAGGAATT AAGAGGATAC ACACGTTGT TTCTCGTGC CTGTTTATG  
 2851 TGCACACATT AGGCATTGAG ACTTCAAGCT TTTCTTTTT TGTCACGTA  
 2901 TCTTTGGTC TTTGATAAAG AAAAGAATCC CTGTTCATG TAAGCACTTT  
 2951 TACGGGGCGG GTGGGGAGGG GTGCTCTGCT GGTCTCAAT TACCAAGAAT  
 3001 TCTCCAAAAC AATTTCTGC AGGATGATTG TACAGAATCA TTGCTTATGA  
 3051 CATGATCGCT TTCTACACTG TATTACATAA ATAAATTAAA TAAAATAACC  
 3101 CCGGGCAAGA CTTTCTTG AAGGATGACT ACAGACATTA AATAATCGAA  
 3151 GTAATTTGG GTGGGGAGAA GAGGCAGATT CAATTTCTT TAACCAGTCT  
 3201 GAAAGTTCAT TTATGATACA AAAGAAGATG AAAATGGAAG TGGCAATATA  
 3251 AGGGGATGAG GAAGGCATGC CTGGACAAAC CCTTCTTTA AGATGTGTCT  
 3301 TCAATTGTA TAAAATGGTG TTTTCATGTA AATAAAATACA TTCTTGGAGG  
 3351 AGC

PRMRD3&gt;



European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number

EP 93 10 5718

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
X	EP-A-0 451 700 (MILES INC.) *Examples 6-8 and 12; Table II; Claims* ---	1-4,8-11	C12N15/12 C12N15/62 C07K15/00 C12N5/10
X	WO-A-9 014 840 (CALIFORNIA BIOTECHNOLOGY INC.) *page 43, line 20 - page 45, line 34; Example 8; Claims* ---	1,4,8-10	
X	WO-A-8 803 951 (CALIFORNIA BIOTECHNOLOGY INC.) *page 41, line 1 - page 43, line 22; page 52, line 1-23; claims* ---	1,4,8-10	
A	WO-A-9 001 540 (CALIFORNIA BIOTECHNOLOGY INC.) *Example II; Claims* ---	1	
D,A	SCIENCE vol. 248, 1990, pages 492 - 495 S.S. SISODIA ET AL.; 'Evidence that beta-amyloid protein in Alzheimer's Disease is not derived by normal processing' *whole document* ---	1	TECHNICAL FIELDS SEARCHED (Int. Cl.5)
P,X	JOURNAL OF BIOLOGICAL CHEMISTRY vol. 267, 1992, pages 25602 - 25608 S.R. SAHASRABUDHE ET AL.; 'Release of amino-terminal fragments from amyloid precursor protein reporter and mutated derivatives in cultured cells' *whole document* -----	1-11	C12N C07K
The present search report has been drawn up for all claims			
Place of search	Date of completion of the search	Examiner	
MUNICH	16 AUGUST 1993	YEATS S.	
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ..... & : member of the same patent family, corresponding document	
X : particularly relevant if taken alone			
Y : particularly relevant if combined with another document of the same category			
A : technological background			
O : non-written disclosure			
P : intermediate document			